



Etude des expressions hors-contexte de gènes tissus-spécifiques dans le cancer du poumon

Aurore Le Bescont Le Bescont

► To cite this version:

Aurore Le Bescont Le Bescont. Etude des expressions hors-contexte de gènes tissus-spécifiques dans le cancer du poumon. Sciences agricoles. Université de Grenoble, 2013. Français. NNT : 2013GRENV068 . tel-01062177

HAL Id: tel-01062177

<https://theses.hal.science/tel-01062177>

Submitted on 9 Sep 2014

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

THÈSE

Pour obtenir le grade de

DOCTEUR DE L'UNIVERSITÉ DE GRENOBLE

Spécialité : **Biologie du développement - Oncogénèse**

Arrêté ministériel : 7 août 2006

Présentée par

Aurore LE BESCONT

Thèse dirigée par **Sophie Rousseaux**

préparée au sein du **Centre de Recherche Oncologie/Développement**
Institut Albert Bonniot
CRI INSERM/UJF U823
dans l'**École Doctorale Chimie et Sciences du Vivant**

Etude des expressions hors-contexte de gènes tissus-spécifiques dans le cancer du poumon

Thèse soutenue publiquement le **12 juillet 2013**,
devant le jury composé de :

Professeur Elisabeth BRAMBILLA

Université Joseph Fourier – Grenoble 1

Présidente

Docteur Sophie POLO

Université Diderot – Paris 7

Rapporteur

Docteur Robert DANTE

Université Claude Bernard – Lyon 1

Rapporteur

Docteur Sophie ROUSSEAUX

Université Joseph Fourier – Grenoble 1

Directrice de thèse



Ahae! mon cœur est lourd de tout ce qui m'échappe.

Le Livre de la Jungle, Rudyard Kipling

Remerciements

J'adresse mes plus vifs remerciements aux membres du jury de cette thèse

le **Professeur Elisabeth Brambilla** pour assurer la présidence de ce jury

les **Docteurs Sophie Polo et Robert Dante** pour avoir accepté la lourde charge de rapporteur.

Soyez assurés de mon respect et de ma reconnaissance pour l'intérêt que vous portez à ce travail.

Ils ont guidé mes pas durant ces trois années, je souhaite remercier

le **Docteur Sophie Rousseaux** pour la confiance qu'elle m'a accordée et ses encouragements précieux

le **Docteur Saadi Khochbin** pour sa disponibilité en toute occasion et ses conseils avisés.

Vous m'avez accompagnée et encouragée dans la voie de la recherche, mais également soutenue dans mes aspirations professionnelles et lors de cette fin de thèse haletante, je vous en suis très reconnaissante.

les **Docteurs Odile Filhol-Cochet et Nadia Alfaidy-Benharouga** pour leur participation à mon comité de suivi de thèse.

Je remercie nos collaborateurs sur ce projet

les Professeurs **Elisabeth et Christian Brambilla** (IAB – CHU Grenoble) pour la cohorte de patients

le Docteur **Philippe Guardiola et Diane Lambert** (Angers) pour les analyses transcriptomiques

le Docteur **Nadia Alfaidy-Benharouga** (CEA, Grenoble) pour les échantillons d'ARN placentaire

nos collègues japonais, les Professeurs **Minoru Yoshida, Kunio Shiota et Mari Shimura**, ainsi que **Saori, Aki et Matzu** pour leur gentillesse et leur prévenance lors de notre séjour à Tokyo

Nicolas Lemaître (CHU, IBP), **Jacques Mazzega** (Microscopie) et **Mylène Pézet** (Cytométrie) pour leur accompagnement technique.

Rien n'aurait été pareil sans la brochette de filles de l'équipe 6

Emilie, prêtresse de la chromatine, dont la douceur et les petites chèvres ont su me rendre le sourire

Anne-Laure, reine de qPCR, qui m'a formée à mon arrivée et a gardé un œil bienveillant sur ce projet

Alexandra, ingénieure et qui le vaut bien, dont les compétences bio-informatiques ont été précieuses

Sandrine, formatrice FACS et gentille fée du labo, dont il n'y a vraiment aucune raison d'avoir peur

Jessica, plus vraiment de l'équipe 6, mais qui m'a secondé audacieusement pendant quelques mois

et tous les autres membres de l'équipe 6, plus particulièrement

Thierry Buchou, pêcheur de bar, qui m'a initiée à la technique du Northern blot avec passion

Fayçal Bousouar pour ses conseils de Bio Mol et nos discussions enflammées à propos d'étudiants

Karin Sadoul, fréquente acolyte de culture cellulaire

ainsi que **Michel Pabion, Sophie Barral, Mahya et Hitoshi**.

Je n'oublie pas les rencontres au sein de l'IAB

Jonathan, Gaëtan, Clothilde, Boubou, Sam, Lydia, Christine, Nicolas, Cyril, Eric, Aude...

Je souhaite remercier tous ceux qui m'ont accompagnée dans ma tâche d'enseignement au sein des équipes pédagogiques de l'UFR de Biologie de l'Université Joseph Fourier, et plus particulièrement :

Gabrielle Tichtinsky, Virginie Stoppin-Mellet, Stefan Nonchev, Ludovic Pelosi, Annie Ray, Virginie

Faure, Sylviane Hennebicq, Sakina Torch, Corinne Mercier et Géraldine Pawlak

pour m'avoir transmis, chacun à leur manière, leur amour de l'enseignement

Catherine Bougault et **Claire Vourc'h** pour m'avoir fait confiance et m'avoir permis d'encadrer une étudiante de *Boston University*, **Ayia A. Mohammed**, durant son semestre en France

Les équipes du **DFI** (Service Doctoral pour la Formation, l'Initiation et l'insertion professionnelles) et du **SUP** (Service Universitaire de Pédagogie) de l'Université de Grenoble pour les formations dispensées

Les **étudiants** que j'ai suivis, puissent vos parcours être riches et passionnants !

Je souhaite adresser une pensée à mes anciens professeurs de S.V.T. qui m'ont donné le goût de l'enseignement, des sciences et de l'enseignement des sciences

Madame Le Borgne (Collège François Truffaut – Betton), **Madame Lesne** (Lycée Jean Macé - Rennes), **Messieurs Poulizac** et **Duprez** (Classe préparatoire BCPST – Lycée Châteaubriand - Rennes), **Déborah Prévot**, **David Busti**, **Vincent Lignier**, **Sandrine Heusser** et **Tristan Ferroir** (Ecole Normale Supérieure - Lyon).

J'espère mettre dans mes cours autant d'énergie / pédagogie / passion / rigueur / humour que vous !

Une pensée émue à mes copains de théâtre (rencontrés au **Créarc**) et à nos aventures sur les planches aux habitués du **TER** Aix-Grenoble
au restaurant de Sushi d'Aix-Les-Bains, à ma paire de skis et à mon VTT
à cette région qui nous a émerveillés et que nous ne voudrions pas quitter.

Je souhaite remercier **mes amis**, que j'ai souvent délaissés dernièrement, mais qui mettent du soleil dans la vie
Aurélia, Emilie, Morgane, Camille, les Lucies, Céline, Laura, Valérie - et leurs moitiés
Arnaud, Fillau, Ludivine, Gildas, Maël, David, Sophie, Max et Olivier, Laure, Charlène...

J'embrasse particulièrement mes deux témoins, **Lydie** et **Valentine**.

Vous êtes à mes côtés à chaque étape de ma vie et j'espère être toujours là pour vous.

Pour finir, je souhaite embrasser du fond du cœur mes proches

Olive : ha ! qu'il est bon d'être un chat !

ma belle-famille : Françoise, Anne-Laure, François, Jean-Paul et Michèle

mes mamies : mes pensées ont été vers vous chaque jour de la rédaction de ce manuscrit

ma sœur Coralie et **mon frère Aymeric** : est-il possible de se manquer autant ? J'espère que nous pourrons toujours sauter dans des trains ou traverser des océans pour nous retrouver. Je vous aime !

mes parents : pour ce que vous avez fait, ce que vous faites et ce que vous ferez pour moi. Que ces pages (peut-être davantage ce qu'elles représentent que leur véritable contenu, qui restera peut-être un peu obscur à vos yeux), soient à la hauteur des sacrifices, du soutien, de la confiance et des encouragements dont vous faites preuve depuis toujours... Je vous aime !

mon Pierre, mon pilou, mon ours gentil et futur mari : durant le mois qui s'est écoulé, tu as été mon roc, mon coach, mon cuisinier. Merci pour tes encouragements, tes conseils, ta patience et ta douceur. Tout est tellement plus facile avec toi à mes côtés. Je t'aime, miaou.

Résumé

Chaque cellule du corps humain possède un génome portant l'ensemble de l'information génétique nécessaire à la constitution de l'organisme entier. Néanmoins, les cellules différenciées n'expriment qu'un répertoire restreint de gènes indispensables au fonctionnement de la cellule. Le contrôle des expressions des gènes est donc fondamental dans la mise en place et le maintien de l'identité cellulaire. Le premier niveau de régulation de l'expression des gènes, le contrôle transcriptionnel, est basé sur l'intégrité de la séquence génique d'une part et sur son accessibilité au sein de la chromatine d'autre part, elle-même conditionnée par un ensemble de mécanismes épigénétiques contrôlant la dynamique chromatinienne. Dans un contexte pathologique, des altérations génétiques et/ou épigénétiques pourront alors être à l'origine de dérégulations géniques et mener à des fonctions cellulaires altérées ou aberrantes.

Lors du développement d'un carcinome bronchique, les cellules cancéreuses pulmonaires acquièrent une capacité de prolifération incontrôlée et une résistance accrue à l'apoptose. Ces caractéristiques phénotypiques, bases de la croissance tumorale, sont le reflet des anomalies qui s'accumulent dans le génome des cellules cancéreuses. Aux altérations génétiques somatiques telles que les mutations ponctuelles ou les remaniements chromosomiques, s'ajoute un bouleversement global du paysage épigénétique, le tout conduisant une crise de l'identité cellulaire et à des dérégulations massives des expressions géniques. Alors que le phénomène de répression aberrante de gènes (notamment des gènes suppresseurs de tumeur) a été largement étudié, l'activation ectopique de gènes normalement silencieux demeure un aspect plus méconnu. Notre hypothèse de travail est que les expressions hors-contexte de gènes tissu-spécifiques, dont les gènes spécifiques de la lignée germinale mâle, puissent être non seulement impliquées dans le phénomène de cancérogenèse, mais également constituer de potentiels biomarqueurs de tumeur ou cibles thérapeutiques innovantes.

Dans cette thèse, notre attention s'est portée sur le gène *PRL* codant la prolactine, normalement exprimée principalement dans l'hypophyse et absente du poumon non tumoral. Nous avons détecté une activation ectopique du gène *PRL* dans 10% des tumeurs pulmonaires, principalement des tumeurs neuroendocrines. Nous avons observé que l'expression de *PRL* est associée à des tumeurs particulièrement agressives et à un pronostic sombre pour les patients. Nous avons également démontré que l'expression de *PRL* confère aux cellules cancéreuses pulmonaires une résistance accrue à un stress génotoxique. De manière inattendue, nos résultats suggèrent que l'effet oncogène de l'expression de *PRL* ne repose pas sur les mécanismes d'action classique de la prolactine et nous avons dû remettre en question l'hypothèse initiale d'une sécrétion de l'hormone par les cellules cancéreuses pulmonaires et d'une action autocrine/paracrine au sein de la tumeur via l'activation du récepteur à la prolactine. Le récepteur est en effet absent des cellules cancéreuses pulmonaires et le transcrit *PRL* exprimé est dépourvu de ses premiers exons, ce qui pourrait conduire à la production d'une protéine tronquée de son peptide signal, incapable d'emprunter la voie de sécrétion classique, et par conséquent retenue à l'intérieur de la cellule cancéreuse pulmonaire. Même si les mécanismes d'action *in cellulo* de la prolactine restent à décrypter, nos données suggèrent que l'expression ectopique de *PRL* pourrait constituer une nouvelle cible thérapeutique dans le traitement des tumeurs pulmonaires agressives.

Ce travail de thèse, incluant également des résultats complémentaires obtenus sur trois gènes spécifiques du testicule exprimés de manière aberrante dans les tumeurs pulmonaires (*BRDT*, *SOX3o* et *SPATA22*) met en lumière l'intérêt de l'étude des expressions ectopiques. Celles-ci peuvent en effet fournir de nouveaux outils de diagnostic et de pronostic aux cliniciens, mais également de nouvelles approches ciblées en complément des thérapies classiques qui ne suffisent pas à limiter la mortalité importante due aux néoplasies pulmonaires.

Mots-clés : cancer du poumon, expressions ectopiques, gène Cancer/Testis, prolactine, oncogène, inhibiteur de HDAC, biomarqueurs de tumeur, biomarqueurs de pronostic.

Abstract

Each human cell contains a genome carrying all the genetic information necessary for the constitution of the whole organism. However, differentiated cells express only a restricted repertoire of genes necessary for cell function. The control of gene expressions is fundamental for the establishment and maintenance of cell identity. The first level of gene expression regulation, the transcriptional control, is based on the integrity of the gene sequence but also on its accessibility within the chromatin, itself dependent on a set of epigenetic mechanisms that control chromatin dynamics. In a pathological context, genetic and epigenetic alterations can lead to gene deregulations and altered or aberrant cellular functions.

During the development of a bronchial carcinoma, lung cancer cells acquire a capacity of uncontrolled proliferation and an increased resistance to cell death. These phenotypic characteristics, favoring tumor growth, result from abnormalities that accumulate in the genome of cancer cells. These are somatic genetic alterations, from point mutations to large-scale chromosomal rearrangements, but also a global disruption of the epigenetic landscape – both leading to an “identity crisis” and to gene deregulations. While the phenomenon of aberrant gene repression (including repression of tumor suppressor genes) has been extensively studied, ectopic activation of normally silent genes remains poorly understood. Our hypothesis is that the “out of context” expression of tissue-specific genes, and more particularly the activation in somatic cancer cells of male germ line-specific genes, not only could be involved in carcinogenesis, but could also be of high interest as tumor biomarkers or novel therapeutic targets.

In this work, we focused on the prolactin-encoding *PRL* gene, normally mainly expressed in the pituitary gland and absent from non-tumor lung. We detected an ectopic *PRL* gene activation in 10% of lung tumors, mainly neuroendocrine tumors. We observed that the expression of *PRL* is associated with aggressive tumors and a poor prognosis for patients. We also found that the expression of *PRL* is associated with an increased resistance of lung cancer cells to a genotoxic stress. Unexpectedly, our data suggest that the oncogenic action of *PRL* expression is not based on the conventional mechanisms of prolactin action, and we did not confirm the initial hypothesis of a secretion by lung cancer cells of the prolactin hormone, and its action in an autocrine / paracrine loop within the tumor through the activation of the prolactin receptor. Indeed, the receptor is absent in lung cancer cells and the transcribed *PRL* mRNA is missing its first exons, possibly leading to the production of a truncated prolactin protein, without a functional signal peptide, therefore unable to follow the classical secretion pathway and retained inside the cancer cell, acting *in cellulo*. Although the detailed mechanisms of prolactin action in lung cancer remain to be deciphered, our study suggests that the ectopic expression of *PRL* could be used as a new therapeutic target in the treatment of aggressive lung tumors.

This work, also including additional results on three testis-specific genes aberrantly expressed in lung tumors (*BRDT*, *SOX30* and *SPATA22*) highlights the interest of studying ectopic gene expressions in tumor cells, which can provide new diagnosis and prognosis tools for clinicians as well as new targeted approaches that could be used in addition to conventional lung cancer therapies, which are presently insufficient to limit the high mortality due to lung neoplasms.

Keywords: lung cancer, ectopic expressions, CT gene, prolactin, oncogene, HDAC inhibitor, tumor biomarker, prognosis biomarker

Avant-propos

Cette thèse a été réalisée dans l'équipe « Epigénétique et signalisation cellulaire » dirigée par Dr. Saadi Khochbin, au sein de l'Institut Albert Bonniot (CRI INSERM/UJF U823) de Grenoble. Les travaux de thèse présentés ci-après ont été dirigés par Dr. Sophie Rousseaux.

Ce doctorat a été conditionné par l'obtention d'une allocation spécifique Normaliens auprès du Ministère de l'Enseignement Supérieur et de la Recherche. Il a réalisé sous un contrat doctoral avec mission d'enseignement de l'Université Joseph Fourier du 01/09/2010 au 31/08/2013.

Table des matières

Remerciements	6
Résumé	8
Abstract	9
Avant-propos	10
Table des matières	11
Tables des figures	17
Liste des abréviations	19

INTRODUCTION 24

1. LA REGULATION DE L'EXPRESSION DES GENES 26

1.1. Mécanismes d'expression génique 26

1.1.1. Structure d'un gène 26

1.1.1.1. Locus 26

1.1.1.2. Unité de transcription 28

1.1.1.3. Promoteur et séquences régulatrices 28

1.1.2. La transcription, étape fondamentale de l'expression d'un gène 30

1.1.3. Le paysage épigénétique contrôle l'accessibilité de la chromatine 32

1.1.3.1. Le nucléosome, unité structurale fondamentale de la chromatine 32

1.1.3.2. Dynamique de l'organisation chromatinienne 32

1.1.3.3. Marques et acteurs épigénétiques 34

1.1.3.3.1. La méthylation de l'ADN 34

1.1.3.3.2. Les modifications des histones 36

1.1.3.3.3. Les complexes de remodelage de la chromatine 38

1.2. Mécanismes de régulation des expressions géniques 38

1.2.1. Mécanismes d'activation transcriptionnelle 40

1.2.1.1. Facteurs de transcription spécifiques à activité activatrice 40

1.2.1.2. Mécanismes activateurs liés à l'acétylation locale des histones 42

1.2.2. Mécanismes de répression transcriptionnelle 42

1.2.2.1. Facteurs de transcription spécifiques à activité répressive 42

1.2.2.2. Mécanismes répresseurs liés à la méthylation des histones 44

1.2.2.3. Mécanismes répresseurs liés à la méthylation locale de l'ADN 44

1.3. Répertoire restreint d'expressions géniques et identité cellulaire 46

2. LE CANCER DU POUMON 50

2.1. Epidémiologie et étiologie du cancer du poumon 50

2.1.1. Un cancer fréquent et de pronostic sombre 50

2.1.1.1. Incidence : Nombre de nouveaux cas détectés chaque année 50

2.1.1.2. Pronostic : Données de survie à 5 ans 50

2.1.1.3. Mortalité : Nombre de décès par an 52

2.1.2. Un cancer imputable en majorité à la consommation de tabac 52

2.1.2.1. Le tabac et les carcinogènes de la fumée de cigarette 52

2.1.2.2. Autres facteurs environnementaux carcinogènes 54

2.1.2.3. Faible contribution de facteurs génétiques familiaux 56

2.2. Carcinomes pulmonaires	56
2.2.1. Développement d'un carcinome pulmonaire	56
2.2.1.1. La muqueuse bronchique non pathologique et ses types cellulaires	56
2.2.1.2. Aspects histologiques de la carcinogenèse pulmonaire	58
2.2.1.3. La stroma-réaction	58
2.2.2. Classification des cancers du poumon	58
2.2.2.1. Classification histopathologique des tumeurs pulmonaires	60
2.2.2.1.1. <i>Les carcinomes bronchiques non à petites cellules CBNPC</i>	60
2.2.2.1.2. <i>Les carcinomes bronchiques à petites cellules CBPC</i>	62
2.2.2.1.3. <i>Des origines cellulaires différentes ?</i>	62
2.2.2.2. Classification selon le stade	62
2.3. Carcinogenèse pulmonaire	64
2.3.1. Altérations de facteurs cellulaires clés : oncogènes et suppresseurs de tumeur	64
2.3.1.1. Activations d'oncogènes	64
2.3.1.2. Inactivations génétiques de gènes suppresseurs de tumeur	66
2.3.2. Détournement de fonctions cellulaires et voies cellulaires impliquées	68
2.3.2.1. Altérations moléculaires promouvant la prolifération	70
2.3.2.2. Altérations moléculaires inhibant la mort cellulaire	70
2.4. Prise en charge et traitements des tumeurs pulmonaires	72
2.4.1. Prise en charge initiale : Détection et diagnostic	72
2.4.2. Les traitements conventionnels et leurs limites	74
2.4.2.1. Chirurgie	74
2.4.2.2. Radiothérapie	74
2.4.2.3. Chimiothérapie	75
2.4.3. Traitements émergents : les thérapies ciblées	75
3. LES ALTERATIONS DES EXPRESSIONS GENIQUES DANS LES CELLULES CANCEREUSES PULMONAIRES	78
3.1. Altérations génétiques et épigénétiques globales dans les cellules cancéreuses	78
3.1.1. Vers l'accumulation progressive d'altérations génétiques	78
3.1.1.1. Mutations ponctuelles : théorie de la mutagenèse somatique	78
3.1.1.2. Mutabilité accrue et phénotype d'instabilité génétique	79
3.1.1.3. Instabilité chromosomique et altérations de grande ampleur	80
3.1.2. Les altérations épigénétiques	82
3.1.2.1. Défauts de méthylation de l'ADN	82
3.1.2.1.1. <i>Hyperméthylations localisées</i>	83
3.1.2.1.2. <i>Hypométhylation globale</i>	83
3.1.2.2. Défauts de modifications des histones	84
3.2. Dérégulations des expressions géniques	86
3.2.1. Inactivations aberrantes de gènes suppresseurs de tumeur	86
3.2.1.1. Pertes physiques des loci suppresseurs de tumeurs	86
3.2.1.2. Répressions aberrantes des expressions de suppresseurs de tumeur	86
3.2.1.3. Vers des thérapies épigénétiques ?	87
3.2.2. Surexpressions d'oncogènes	88
3.3. Expressions ectopiques de gènes tissu-spécifiques	90
3.3.1. Historique : la découverte des gènes CT comme cibles d'immunothérapies	90
3.3.2. Ré-expression des facteurs CT dans les cancers pulmonaires	91
3.3.3. Des facteurs CT comme moteurs de la carcinogenèse pulmonaire	92
3.3.4. Un concept étendu aux facteurs tissu-spécifiques	95

RESULTATS**100****CHAPITRE 1: Expression aberrante du gène hypophysaire *PRL* dans les tumeurs pulmonaires neuroendocrines****102****1. CONTEXTE DE L'ETUDE****104****1.1. La prolactine, hormone peptidique hypophysaire****104**

1.1.1. Structure de la protéine

104

1.1.2. Sécrétion par les cellules hypophysaires

104

1.2. Le rôle lactogénique et mammotrope de la prolactine**106**

1.2.1. Le récepteur de la prolactine

106

1.2.2. La voie de signalisation lactogénique

106

1.3. D'autres cibles et d'autres sources de la prolactine**107**

1.3.1. Les cibles secondaires de l'hormone hypophysaire

107

1.3.2. Les sources extra-hypophysaires de prolactine

108

2. RESULTATS**110****2.1. Le gène hypophysaire *PRL* est exprimé de manière aberrante par les cellules cancéreuses pulmonaires****110**2.1.1. Le gène *PRL* est identifié dans la signature transcriptomique de tumeurs pulmonaires agressives

110

2.1.2. L'expression du gène *PRL*, restreinte à l'hypophyse et quelques sources extra-hypophysaires, n'est pas détectée dans le poumon sain

112

2.1.3. Le gène *PRL* est exprimé de manière aberrante dans 10% des tumeurs pulmonaires

112

2.1.4. Le gène *PRL* est exprimé de manière aberrante par des lignées de cancer du poumon à petites cellules

116

2.2. L'expression du gène *PRL* dans les cellules cancéreuses pulmonaires est associée à un pronostic sombre et à une chimiorésistance accrue**116**2.2.1. Les tumeurs qui expriment le gène *PRL* sont principalement des tumeurs neuroendocrines

116

2.2.2. L'expression aberrante du gène *PRL* est associée à un pronostic sombre pour le patient

116

2.2.3. L'expression aberrante du gène *PRL* confère aux cellules cancéreuses pulmonaires une plus grande résistance à un stress génotoxique

118

2.3. L'action de la prolactine est indépendante de son récepteur**122**2.3.1. Le récepteur à la prolactine n'est exprimé ni dans le poumon non-tumoral, ni dans les cellules cancéreuses qui expriment le gène *PRL*

122

2.3.2. La prolactine exprimée de manière ectopique pourrait agir *in cellulo* sans être sécrétée

124

2.4. L'expression aberrante du gène *PRL* dans les cellules pulmonaires cancéreuses aboutit à la production d'un ARN messager tronqué en 5'**128**2.4.1. L'expression ectopique du gène *PRL* ne semble due ni à une amplification génique, ni à une déméthylation du promoteur

128

2.4.2. L'expression du gène *PRL* dans les cellules cancéreuses pulmonaires ne repose sur l'utilisation d'aucun des deux promoteurs connus de *PRL*

130

2.4.3. Un ARN messenger tronqué ne contenant pas la séquence de l'exon 1 est produit dans les lignées de cancer du poumon à petites cellules	130
2.4.4. Les ARN messagers synthétisés au sein des tumeurs pulmonaires de patients ne contiennent pas les séquences exoniques en 5'	134
2.5. Quelques données sur l'action <i>in cellulo</i> de la prolactine dans les cellules cancéreuses pulmonaires	134
2.5.1. Mise au point d'une lignée modèle de cancer du poumon exprimant une prolactine non sécrétable	134
2.5.2. La prolactine non sécrétable est clivée dans notre lignée modèle de cancer du poumon	136
2.5.3. L'acétylation de la prolactine non sécrétable contrôle sa localisation subcellulaire	140
2.5.4. La prolactine non-sécrétable pourrait être liée aux voies cellulaires d'acétylation, mais pas via une interaction directe avec l'acétyltransférase p300	142
2.6. Les analyses haut débit des signatures géniques associées à l'expression du gène <i>PRL</i> pourraient aider à décrypter son mode d'action	144
2.6.1. L'expression de <i>PRL</i> régule un large groupe de gènes dans les cellules de CBPC	144
2.6.2. Des cibles potentielles de l'expression de <i>PRL</i> sont identifiées dans les transcriptomes de patients	146
2.7. L'expression aberrante du gène <i>PRL</i> peut être ciblée par des drogues épigénétiques	154
3. DISCUSSION	156
3.1. La nécessité de nouvelles approches thérapeutiques dans la lutte contre les cancers pulmonaires	156
3.2. L'expression ectopique du gène <i>PRL</i> dans les cellules pulmonaires cancéreuses	160
3.3. Prolactine et cancers : une histoire à deux visages	163
3.4. Les cellules pulmonaires cancéreuses produisent un variant d'ARN messenger <i>PRL</i> amputé des premiers exons	166
3.5. La résistance <i>PRL</i> -associée	170
3.6. Hypothèses sur la protéine	171
3.7. Vers l'utilisation de l'expression de <i>PRL</i> en tant que biomarqueur ou cible thérapeutique	172
Chapitre 2: Expression aberrante du facteur testiculaire <i>BRDT</i> dans les tumeurs pulmonaires	176
1. CONTEXTE DE L'ETUDE	178
2. RESULTATS	182
2.1. La protéine <i>BRDT</i> spécifique de la lignée germinale mâle est produite ectopiquement dans une lignée de cancer pulmonaire	182
2.2. La présence ectopique de <i>BRDT</i> induit une plus grande sensibilité des cellules cancéreuses à un traitement par la TSA	182
2.3. Etablissement d'une lignée cancéreuse pulmonaire modèle exprimant <i>BRDT</i> de manière stable	184
3. DISCUSSION	188

Chapitre 3 : Etude des profils d'expression des gènes *SOX30* et *SPATA22* exprimés de manière aberrante dans les tumeurs pulmonaires

190

1. CONTEXTE DE L'ETUDE

192

2. RESULTATS

192

2.1. Profil d'expression de *SPATA22* dans les cellules spermatogéniques

194

2.2. Profil d'expression de *SOX30* dans les cellules spermatogéniques

194

3. DISCUSSION

196

CONCLUSION

198

MATERIELS ET METHODES

202

1. MATERIELS

204

1.1. Matériels biologiques

204

1.2. La cohorte de patients "Cancer du poumon"

204

2. METHODES

206

2.1. Méthodes sur cellules et tissus

206

2.1.1. Purification et fractionnement des cellules germinales mâles

206

2.1.2. Culture de lignées cellulaires cancéreuses

206

2.1.2.1. Lignées cellulaires utilisées et conditions de culture

206

2.1.2.2. Ensemencements de cellules cancéreuses

208

2.1.3. Transfections de cellules cancéreuses en culture

208

2.1.3.1. Transfection de siRNA

208

2.1.3.2. Transfection d'ADN plasmidique

208

2.1.4. Traitements des cellules cancéreuses en culture

210

2.1.5. Immunomarquages

210

2.1.5.1. Marquage immunofluorescent sur cellules cancéreuses

210

2.1.5.2. Cytométrie de flux

210

2.1.5.2.1. Analyse de la mortalité cellulaire en cytométrie de flux

210

2.1.5.2.2. Tri d'une population cellulaire en cytométrie de flux

212

2.1.5.3. Immunofluorescence sur cellules testiculaires

212

2.1.5.4. Immunohistochimie sur coupe testiculaire

212

2.2. Méthodes sur acides nucléiques

214

2.2.1. Extraction d'acides nucléiques

214

2.2.1.1. Extraction ARN : Kit Macherey-Nagel

214

2.2.1.2. Extraction ARN : Kit Qiagen

214

2.2.1.3. Extraction ARN au TRizol	216
2.2.2. Analyse des acides nucléiques	216
2.2.2.1. RT	216
2.2.2.2. PCR	216
2.2.2.3. qPCR	216
2.2.2.4. Electrophorèse d'ADN sur gel d'agarose	220
2.2.2.5. Northern blot	220
2.2.3. Clonages	221
2.3. Méthodes sur protéines	222
2.3.1. Extraction des protéines	222
2.3.1.1. En tampon de lyse dénaturant	222
2.3.1.2. En tampon de lyse douce	222
2.3.1.3. Fractionnement protéique cytosoluble-nucléaire insoluble	222
2.3.1.4. Extrait acide des histones	222
2.3.2. Production de protéine recombinante	223
2.3.2.1. Cinétique d'induction	223
2.3.2.2. Extrait protéique soluble	223
2.3.2.3. Purification d'une protéine recombinante tagguée 6-histidines sur billes agarose-nickel	223
2.3.2.4. Purification d'une protéine recombinante fusionnée à la GST sur billes sépharose glutathion	223
2.3.3. Tests <i>in vitro</i> sur protéines	224
2.3.3.1. Expérience de déphosphorylation <i>in vitro</i> des protéines	224
2.3.3.2. Immunoprécipitation	224
2.3.3.3. Test d'interaction <i>in vitro</i> avec protéine purifiée sur billes	224
2.3.3.4. Peptide pull-down	226
2.3.4. Western-blot	226
2.3.4.1. Electrophorèse SDS-Page	226
2.3.4.2. Transfert	226
2.3.4.3. Immunodétection	226
2.4. Méthodes sur données bioinformatiques	227
2.4.1. Analyse des données de survie et des transcriptomes de patients	227
2.4.2. Analyse transcriptomique des lignées CBPC	227

ANNEXES **230**

Annexe 1 : (Wang et al., 2011)	232
Annexe 2 : (Rousseaux, 2013)	238

BIBLIOGRAPHIE **250**

Table des figures

- Figure 1 : Le gène humain *PRL* : locus, organisation et mode d'expression simplifié
- Figure 2 : Séquences régulatrices du gène *PRL* et facteurs de transcription se liant à ces séquences
- Figure 3 : Transcription d'un gène de classe 2
- Figure 4 : La structure chromatinienne
- Figure 5 : Méthylation de l'ADN
- Figure 6 : Modifications post-traductionnelles des histones
- Figure 7 : Complexes de remodelage de la chromatine
- Figure 8 : Marques épigénétiques, états chromatinien et permissivité à la transcription
- Figure 9 : Mécanismes d'activation transcriptionnelle
- Figure 10 : Mécanismes de répression transcriptionnelle
- Figure 11 : Profils d'expression des gènes, stabilité de l'identité cellulaire et adaptabilité
- Figure 12 : Facteurs de transcription et marques épigénétiques établissent les patrons d'expression des gènes
- Figure 13 : Données chiffrées sur l'incidence et la mortalité des cancers du poumon en France et dans le monde
- Figure 14 : Evolutions de l'incidence et de la mortalité du cancer du poumon durant les dernières années
- Figure 15 : Implication de la consommation de tabac dans les cancers pulmonaires
- Figure 16 : Carcinogènes du tabac et cancers pulmonaires
- Figure 17 : L'épithélium bronchique normal et pathologique
- Figure 18 : Classification histopathologique simplifiée des tumeurs pulmonaires
- Figure 19 : Grille de classification TNM des tumeurs pulmonaires
- Figure 20 : Grille de classification par stades des tumeurs pulmonaires
- Figure 21 : Mutations de l'oncogène KRAS et du suppresseur de tumeur p53 dans les cancers pulmonaires
- Figure 22 : Les caractéristiques fondamentales des cellules cancéreuses selon Hanahan et Weinberg
- Figure 23 : Voies de signalisation couramment altérées dans les cellules cancéreuses pulmonaires
- Figure 24 : Référentiel de prise en charge et de traitement des cancers bronchiques
- Figure 25 : Thérapies émergentes ciblant des composants des voies de signalisation
- Figure 26 : La compréhension des causes génétiques et épigénétiques dans les cancers
- Figure 27 : Perte de matériel génétique et gènes suppresseurs de tumeur associés
- Figure 28 : Hyperméthylation des promoteurs de gènes suppresseurs de tumeur
- Figure 29 : Principe théorique de la vaccination anticancéreuse par un antigène spécifique de tumeur
- Figure 30 : Activation ectopique de gènes CT dans les cellules cancéreuses pulmonaires
- Figure 31 : Sécrétion de prolactine par les cellules lactotrophes de l'anté-hypophyse
- Figure 32 : Le récepteur PRL-R médie l'effet mammotrope et lactogénique de la prolactine
- Figure 33 : *PRL* gene is identified in the gene signature of aggressive tumors
- Figure 34 : *PRL* gene expression in human tissues: predominant in pituitary gland, some extra-pituitary sources
- Figure 35 : *PRL* gene is expressed within the pulmonary tumor of 9,2% patients
- Figure 36 : Identification of *PRL* expressing-small cell lung cancer cell lines
- Figure 37 : Characterization of *PRL*+ patients
- Figure 38 : *PRL* gene aberrant expression within a tumor is associated with a poor prognosis for the patient
- Figure 39 : Effect of *PRL* silencing on *PRL*+ SCLC cell lines
- Figure 40 : Effect of *PRL* silencing on cell mortality in *PRL*+ SCLC cell lines
- Figure 41 : Expression pattern of *PRL-R* gene in human tissues and lung cancer cell lines
- Figure 42 : Expression pattern of *PRL-R* gene in lung cancer patients
- Figure 43 : PRL protein is neither detected in *PRL*+ SCLC cell lines extracts, nor in cell culture supernatants
- Figure 44 : An impaired secretion of prolactin due to a low functional Golgi apparatus?
- Figure 45 : Mechanisms of *PRL* gene re-expression in lung cancer cells
- Figure 46 : The synthesis of *PRL* mRNA does not rely on proximal or distal promoters

Figure 47 : Exon 1 and exon 1a sequences are absent of the SCLC cell lines *PRL* mRNA

Figure 48 : Analysis of the *PRL* mRNA produced by H146 SCLC *PRL*+ cells by Northern blot

Figure 49 : Differently extended absences of 5' sequences in *PRL* mRNA from *PRL*+ patients tumor

Figure 50 : Expression of secretable PRL-GFP and non-secretable GFP-PRL in the H1299 lung cancer cell line

Figure 51 : Effects of non-secretable GFP-PRL expression in the H1299 human cancer cell line

Figure 52 : Non secretable GFP-PRL is cleaved in a truncated PRL16 form with potential antiproliferative role

Figure 53 : Acetylation of non secretable GFP-PRL controls its subcellular localization

Figure 54 : PRL does not physically interact with the histone acetyltransferase p300

Figure 55 : Differential transcriptome in *PRL*-silenced versus *PRL*-expressing SCLC cells

Figure 56 : Cellular components and biological processes affected by the silencing of *PRL*

Figure 57 : GSEA analysis of differential transcriptomes in SCLC identifies significantly concordant gene sets

Figure 58 : Differential transcriptome in *PRL*-expressing versus *PRL*-negative lung tumors

Figure 59 : Identification of *PRL*-coregulated genes and of *PRL*-/ *PRL*+ like patients

Figure 60 : Comparison of differentially expressed-genes in patients and cell lines

Figure 61 : *PRL* silencing and *PRL*+ SCLC cell mortality induced by the HDAC inhibitor Trichostatin A

Figure 62 : La carcinogénèse pulmonaire, reflet des altérations qui s'accumulent dans le génome

Figure 63 : Patrons d'altérations moléculaires distincts entre CBPC et CBNPC

Figure 64 : Fréquences de l'activation ectopique du gène *PRL* dans les tumeurs pulmonaires de la cohorte

Figure 65 : Prolactine et cancer

Figure 66 : Expliquer l'absence des premiers exons dans l'ARNm *PRL* produit dans le cancer du poumon

Figure 67 : Autres débuts de traduction et autres cadres de lecture potentiels pour l'ARNm *PRL*

Figure 68 : Vers un mode d'action complètement inédit d'une prolactine tronquée non sécrétable

Figure 69 : BRDT, un facteur à bromodomains spécifique de la lignée germinale mâle

Figure 70 : BRDT est impliqué dans le remodelage de la chromatine acétylée

Figure 71 : BRDT, facteur spécifique du testicule, est exprimé dans les cellules cancéreuses pulmonaires

Figure 72 : BRDT est impliqué dans la mortalité induite par la TSA des cellules cancéreuses pulmonaires H827

Figure 73 : Caractérisation des lignées stables H1299 exprimant des formes de BRDT fusionnées à la GFP

Figure 74 : La TSA induit une compaction anormale de la chromatine dans le clone H1299 BRDT Δ C-GFP

Figure 75 : *SOX30* et *SPATA22*, gènes testiculaires, sont activés dans les tumeurs pulmonaires

Figure 76 : Expression pattern of *SPATA22* during male germ cell differentiation in rodent testes

Figure 77 : Expression pattern of *SOX30* during male germ cell differentiation in mouse testes

Figure 78 : Lignées cellulaires utilisées : provenance, caractéristiques et conditions de culture.

Figure 79 : Conditions d'ensemencement de lignées cellulaires cancéreuses selon l'utilisation ultérieure

Figure 80 : Protocoles de transfection des cellules cancéreuses

Figure 81 : Traitements des cellules cancéreuses en culture

Figure 82 : Aperçu de la méthode d'analyse par cytométrie en flux

Figure 83 : Listes des anticorps utilisés et conditions d'utilisation

Figure 84 : Liste des primers utilisés

Figure 85 : Localisation des différents primers et siRNA dans la séquence des variants ARNm *PRL*

Figure 86 : Aperçu de la méthode d'analyse par qPCR

Figure 87 : Aperçu du protocole de clonage

Figure 88 : Composition des gels de concentration en fonction de la taille de la protéine d'intérêt

Figure 89 : Echantillons d'ARN des lignées de CBPC H146 et H524 envoyés pour l'analyse transcriptomique

Listes des abréviations et acronymes

ABC	Avidin Biotin Complex
Acetyl-CoA	Acétyl-Coenzyme A
ACTH	Hormone corticotrope
Acti D	Actinomycine D
ADC	Adénocarcinome
ADH	Hormone anti-diurétique
ADN	Acide désoxyribonucléique
ADNc	Acide désoxyribonucléique complémentaire
AFSSET	Agence française de sécurité sanitaire de l'environnement et du travail
AMPC	Adénosine MonoPhosphate cyclique
ANP	Facteur natriurétique auriculaire
AP-1	Activator protein 1
ARC	Association pour la Recherche sur le Cancer
ARN	Acide RiboNucléique
ARNm	Acide RiboNucléique messenger
miARN	micro Acide RiboNucléique
siARN	Petit ARN interférent
ALK	Anaplastic Lymphoma Kinase
ALT	Alternative Lengthening of Telomeres
ATP	Adénosine tri-phosphate
AVP	Arginine VasoPressine
5AzaCdR	5-aza-2'-deoxycytidine
BAC	Chromosome artificiel bactérien - <i>Bacterial Artificial Chromosome</i>
BD	Bromodomaine
BLAST	Basil Local Alignment and Search Tool
BN/GRP	Bombésine / Gastrin releasing peptide
bp	Base pair
BET	Bromure d'éthidium
BET	Bromodomain and Extra-Terminal domain
BRCA1	BReast CAncer 1, early onset
BRDT	BRomoDomain Testis-specific protein
BRE	TFIIB Recognition Element
BSA	Albumine sérique bovine - <i>Bovine Serum Albumine</i>
BTB	Barrière hémato-testiculaire - <i>Blood-Testis Barrier</i>
CBP	CREB Binding Protein
CBNPB	Carcinome bronchique non à petites cellules
CBPB	Carcinome bronchique à petites cellules
CD	Séquence codante
CDK	Kinase cycline-dépendante – <i>Cycline dependent kinase</i>
CDKN2A	Cyclin-dependent kinase inhibitor 2A
CEB	Carcinome épidermoïde bronchique
C/EBP	CAAT Enhancer-Binding Protein
CGH	Hybridation génomique comparative - <i>Comparative Genomic Hybridization</i>
CIM-O	Classification internationale des maladies oncogéniques
CIN	Instabilité chromosomique - <i>Chromosome Instability</i>

CIP	Calf Intestinal alkaline Phosphatase
CIRC	Centre International de Recherche sur le Cancer
CKII	Caséine kinase II
CNV	Variation du nombre de copies - <i>Copy Number Variation</i>
CpG	Dinucléotide Cytosine-Guanosine
CPGC	Carcinome pulmonaire à grandes cellules
CPNEGC	Carcinome pulmonaire neuroendocrine à grandes cellules
CPI	Complexe de Pré-Initiation de la transcription
CRE	cAMP-Responsive Element
CREB	cAMP Response Element Binding protein
Ct	Cycle threshold
CT	Tomodensitométrie – <i>Computed Tomography</i>
CT	Facteur testiculaire du cancer - <i>Cancer/Testis gene</i>
CT-X	Gène CT situé sur le chromosome X
CT-non X	Gène CT non situé sur le chromosome X
CTD	Carboxy-terminal domain
CTP	Cytidine tri-phosphate
DAB	3, 3'-DiAminoBenzidine
DBD	Domaine de liaison à l'ADN – <i>DNA Binding Domain</i>
DEPC	Diéthyl pyrocarbonate
DMEM	Dulbecco's Modified Eagle Medium
DMR	Differentially Methylated Region
DMSO	Diméthylsulfoxyde
DNMT	DNA cytosine-5 methyltransferase
dNTP	Désoxy-Nucléotides Tri-Phosphate
DO	Densité Optique
DSB	Cassure double-brin de l'ADN - <i>DNA Double-Strand Breaks</i>
DSIR	Designer of Small Interfering RNA
DTT	Dithiothréitol
ECD	Domaine extra-cellulaire – <i>Extra-Cellular Domain</i>
EGF	Facteur de croissance épidermique – <i>Epidermic Growth Factor</i>
EGF-R	Récepteur à l'EGF
EML4	Microtubule-associated protein-Like 4
EPO	Erythropoïétine
ES	Cellule souche embryonnaire – <i>Embryonic Stem cell</i>
EST	Expressed Sequence Tag
EtOH	Ethanol
Ets	E-twenty six transcription factor
FACS	Cytomètre de flux
FHIT	Gène fragile histidine triade
FISH	Hybridation <i>in situ</i> en fluorescence - <i>Fluorescent in situ Hybridization</i>
Francim	Réseau français des registres des cancers
FSC	Forward Scatter
FSH	Hormone folliculo-stimulante - <i>Follicle-Stimulatin Hormone</i>
FT	Facteur de transcription
GAS	Gamma-interferon-Activated Sequence
GEO	Gene Expression Omnibus
GFP	Green Fluorescent Protein

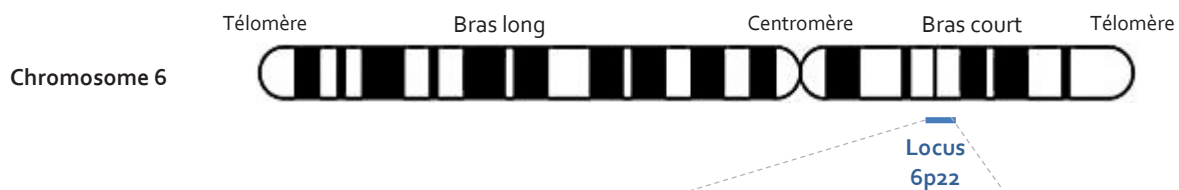
GH	Hormone de croissance – <i>Growth Hormone</i>
GM-CSF	Granulocyte Macrophage-Colony Stimulating Factor
GO	Gene Ontology
GSEA	Gene Sets Enrichment Analysis
GST	Glutathion S-Transférase
GTP	Guanosine-triphosphate
HAP	Hydrocarbures aromatiques polycycliques
HAT	Histone acétyltransférase
HDAC	Histone déacétylase
HDACi	Inhibiteur des histones déacétylases
HDM	Histone déméthylase
HIF	Hypoxia-Inducible Factor
HLH	Hélice-Boucle-Hélice - <i>Helix-Loop-Helix</i>
HMG	High Mobility Group
HMT	Histone méthyltransférase
HP1	Heterochromatin Protein 1
HRP	Péroxydase du raifort - <i>HorseRadish Peroxidase</i>
HSP	Protéines de choc thermique - <i>Heat Shock Proteins</i>
IARC	Association Internationale de Recherche sur le Cancer
ICD	Domaine intra-cellulaire – <i>Intra-Cellular Domain</i>
ICR	Région de contrôle de l’empreinte - <i>Imprinting Control Region</i>
IF	Immunofluorescence
IFN	Interféron
IHC	Immunohistochimie
INCa	Institut National du Cancer
INED	Institut National d’Etudes Démographiques
INPES	Institut National de Prévention et d’Education pour la Santé
InVS	Institut national de Veille Sanitaire
IP	Immunoprécipitation
IPS	Cellule souche pluripotente induite - <i>Induced Pluripotent Stem cell</i>
IPTG	Isopropyl β -D-1-thiogalactopyranoside
IRM	Imagerie par Résonnance Magnétique
JAK-2	Janus kinase 2
JmJ	Jumonji
KAP-1	KRAB-ZFP Associated Protein 1
Kb	Kilobase
kDa	Kilodalton
KO	Knock-Out
KRAB-ZFP	Krüppel Associated Box – Zinc Finger Protein
KRAS	Kirsten Rat Sarcoma viral oncogene homolog
LB	Lysogeny broth medium
LC-LC	Large-Cell Lung Carcinoma
LCNE-LC	Large-Cell Neuroendocrine Lung Carcinoma
LH	Hormone lutéinisante - <i>Luteinizing hormone</i>
LIF	Leukemia Inhibitory Factor
LOH	Perte d’hétérozygotie - <i>Loss of Heterozygoty</i>
LSD1	Lysine Specific Demethylase 1

MAGE	Melanoma antigen gene
MBD	Methyl CpG Binding Domain
MBP	Methyl CpG Binding Protein
MGF	Mammary Gland Factor
MGMT	O6-methylguanine-DNA méthyltransférase
MIQE	Minimum Information for Publication of Quantitative Real-Time PCR Experiments
MM	Masse moléculaire
MOPS	Acide 3-(N-morpholino)propanesulfonique
MPT	Modification Post-Traductionnelle
MW	Molecular Weight
NER	Système de réparation de l'ADN par excision de nucléotides - <i>Nucleotide Excision Repair</i>
NES	Signal d'export nucléaire - <i>Nuclear Export Signal</i>
NK	Lymphocyte <i>Natural Killer</i>
NMT	Carcinome de la lignée médiane - <i>NUT midline carcinoma</i>
NNK	Nicotine-derived Nitrosamine Ketone (4-(Methylnitrosamino)-1-(3-pyridyl)-1-butanone)
NSCLC	Non Small Cell Lung Carcinoma
NuRD	Nucleosome Remodeling and Deacetylase
NUT	NUclear protein in Testis
NY-ESO-1	New York esophageal squamous cell carcinoma-1
OF₁	Oncins France 1
OMS	Organisation Mondiale de la Santé
ORF	Cadre ouvert de lecture - <i>Open Reading Frame</i>
PAS	Coloration Periodic Acid Schiff
pb	Paire de bases
PBS	Tampon salin phosphate – <i>Phosphate buffer saline</i>
(q)PCR	(quantitative) Polymerase Chain Reaction
PD	Pull-down
PE	Phycoérythrine
PFA	Paraformaldéhyde
PIF	Prolactin Inhibiting Factor
PIT-1	POU-homeodomain transcription factor 1
PHD	Plant HomeoDomain
PKA	Protéine Kinase A
PKC	Protéine Kinase C
PL	Hormone placentaire lactogène – <i>Placental lactogen</i>
POB	PyridylOxoBytil
PRL	Prolactine
PRL-R	Récepteur de la prolactine
PSM	Poste de Sécurité Microbiologique
PTM	Post Translational Modification
RARβ	Récepteur bêta à l'acide rétinoïque
RASSF1A	RAS association domain family 1A
Rb	Retinoblastoma
RMA	Algorithme <i>Robust Multi-array Average</i>
RMN	Résonance Magnétique Nucléaire
rpm	Rotation par minute
RPMI	Roswell Park Memorial Institute medium
RT	Transcription inverse - <i>Reverse Transcription</i>

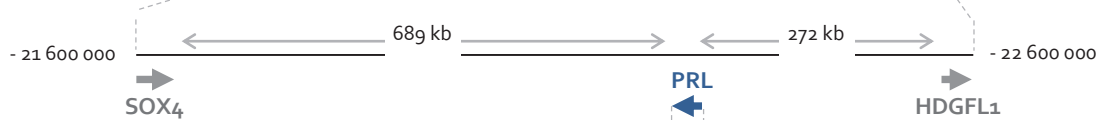
SCLC	Small Cell Lung Carcinoma
SDS	Dodécylsulfate de sodium
SET	Suppressor of the Variegation - Enhancer of Zeste and Trithorax
siRNA	small interfering Ribonucleic Acid
SKY	Caryotype spectral - <i>Spectral Karyotype</i>
SOX	Sex determining Region on Y (SRY)-box
SPATA	Spermatogenesis-associated protein
SQC-LC	Squamous-Cell Lung Carcinoma
SSC	Tampon Saline-Sodium Citrate
SSC	Side Scatter
STAT	Signal Transducers and Activators of Transcription
SVF	Sérum de Veau Foetal
SWI/SNF	mating type Switching/ Sucrose Non Fermenting
TAD	Trans-activating Domain
TAE	Tampon Tris-Acétate-EDTA
TBP	TATA-binding protein
TCA	Acide trichloroacétique
TEP	Tomographie à Emission de Positrons
TF	Transcription factor
TGS	Tampon Tris-Glycine-SDS
TIDA	Neurone dopaminergique tubéroinfundibulaire - <i>Tuberoinfundibular Dopaminergic neurone</i>
TK	Domaine Tyrosine Kinase
TKI	Inhibiteur des domaines Tyrosine Kinase
TRD	Transcriptional Repression Domain
TRE	TPA DNA Response Element
TRH	Hormone thyroïdienne - <i>Thyrotropine Releasing Hormone</i>
TSA	Trichostatine A
TSH	Thyréostimuline - <i>Thyroid-Stimulating Hormone</i>
TSP	Tumor Sequencing Project
TSS	Site d'initiation de la transcription - <i>Transcription start signal</i>
UICC	Union Internationale contre le Cancer
UTR	UnTranslated Region
UV	Ultra-Violet
VEGF	Facteur de croissance de l'endothélium vasculaire - <i>vascular Endothelial Growth Factor</i>
WB	Western blot
WHO	World Health Organization
ZFP	Zinc Finger Protein

INTRODUCTION

A.



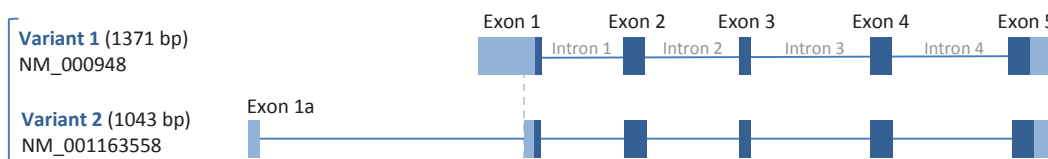
B.



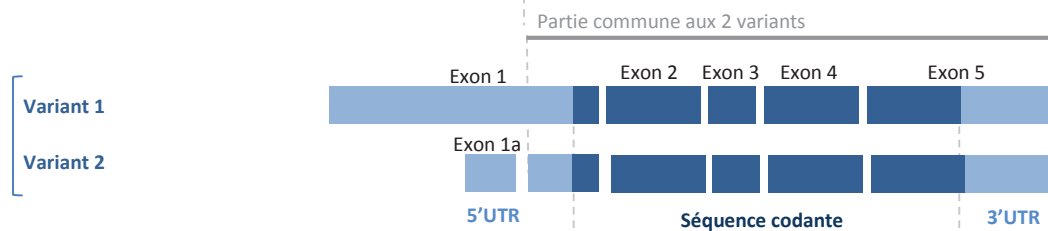
C.



D.



E.



F.

Protéine Prolactine (227 aa)



Figure 1 : Le gène humain *PRL* : locus, organisation et mode d'expression simplifié

- A. Le gène *PRL* est porté par le petit bras du chromosome 6, au niveau du locus 6p22.
- B. Le gène *PRL* est porté par le brin (-) de l'ADN. Il est entouré sur environ 1 mégabase d'un désert de séquences non-codantes.
- C. Le TSS utilisé dans l'hypophyse (TSS2) est précédé 5,3 kb en amont d'un TSS alternatif (TSS1) utilisé dans le placenta.
- D. L'ARNm *PRL* placentaire (variant 2) inclue l'exon 1a non-codant absent du transcrit hypophysaire (variant 1).
- E. Ces deux variants ne diffèrent que dans la partie 5'UTR non codante : la séquence codante est identique.
- F. La même protéine prolactine sera synthétisée lors de la traduction.

1. LA REGULATION DE L'EXPRESSION DES GENES, BASE DE L'IDENTITE ET DE L'HOMEOSTASIE CELLULAIRE

Les expressions géniques, finement régulées, vont permettre l'établissement et le maintien de l'identité cellulaire. A l'inverse, les dérégulations d'expressions géniques observées dans les cancers aboutissent à une crise d'identité et participent au processus de transformation cellulaire. Avant de nous intéresser au modèle du cancer du poumon (**voir partie 2**) et aux altérations d'expressions géniques dans les cellules cancéreuses pulmonaires (**voir partie 3**), cette première partie introductive vise à présenter succinctement les mécanismes qui contrôlent l'expression des gènes dans les cellules saines.

1.1. Mécanismes d'expression génique : comment un gène s'exprime-t-il ?

Le dogme central de la biologie moléculaire (de l'ADN à l'ARN messager puis à la protéine), bien que modifié et complété sur de nombreux aspects au cours des dernières décennies (variants ARN issus des épissages, existence de gènes codant des ARN non messagers, importance des modifications post-traductionnelles des protéines...), énonce le principe fondamental de l'expression d'un gène, selon lequel l'information héréditaire codée dans l'ADN aboutit à la fabrication de molécules actives dans le fonctionnement cellulaire. Ainsi, un gène est dit exprimé lorsqu'une molécule à rôle actif, ARN ou protéine, est produite. Pour un gène codant une protéine (gène de classe 2), ceci nécessite la conjonction de différentes étapes : transcription de la séquence génique en ARN pré-messager, maturation de l'ARN messager (polyadénylation, épissage, coiffage, export nucléaire), synthèse du polypeptide par la machinerie de traduction et/ou modifications post-traductionnelles de la protéine (repliement, clivage protéique, adjonction de groupements chimiques...). Par abus de langage, et notamment lors des analyses basées sur la détection des ARNs (Northern blot, RT-qPCR, analyses transcriptomiques sur puces, RNAseq...), on dit couramment qu'un gène est exprimé dès lors que ce gène est transcrit et que l'ARN messager est produit. Nous allons ici nous attacher à décrire la structure générale d'un gène humain ainsi que les phénomènes aboutissant à la première étape de son expression, la transcription en ARN pré-messager. Etant donnée la variabilité des gènes humains, ceci sera fait non pas dans un souci d'exhaustivité, mais en s'appuyant sur des exemples choisis et notamment celui du gène *PRL* (codant la prolactine).

1.1.1. Structure d'un gène

1.1.1.1. Locus

Un gène peut être défini comme une unité d'information génétique, codée par la séquence de nucléotides de la molécule d'ADN, déterminant un caractère héréditaire. Chaque gène possède son propre locus, emplacement physique précis et invariable sur un chromosome. Ainsi, le gène humain *PRL* est porté par le brin (-) de l'ADN génomique, sur le petit bras du chromosome 6, au niveau du locus 6p22 (**Figure 1A**). La cartographie du locus *PRL* a été réalisée dans les années 1980, d'abord par des analyses en Southern-blot ayant identifié le chromosome 6 comme porteur du gène ([Owerbach et al., 1981](#)), puis par l'analyse de cellules somatiques hybrides souris-humain ayant permis de restreindre l'intervalle du locus au petit bras 6pter-p21.1 ([Taggart et al., 1987](#)) et enfin par la création de délétions du chromosome 6 par gamma-irradiation qui ont établi la localisation du gène *PRL* dans l'intervalle 6p22.2-p21.3 ([Evans et al., 1989](#)). Le gène *PRL* est entouré, sur une mégabase d'ADN, d'un désert de séquences non-codantes hautement conservées ([Semprini et al., 2009](#)) (**Figure 1B**).

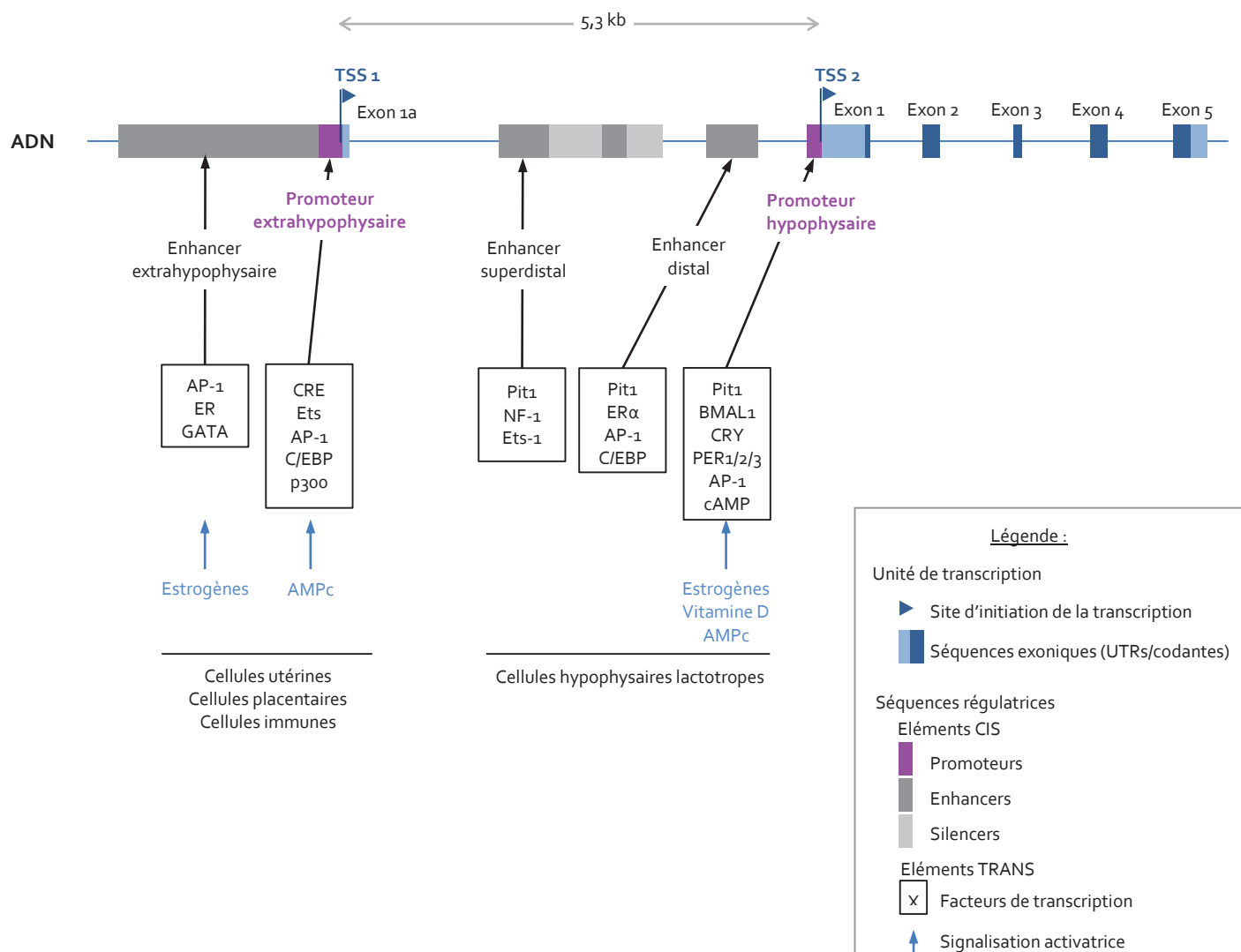


Figure 2 : Séquences régulatrices du gène *PRL* et facteurs de transcription se liant à ces séquences

Adapté de (Featherstone et al., 2012).

Le gène *PRL* possède deux promoteurs alternatifs tissu-spécifiques. Le premier promoteur, situé en amont du TSS2 et de l'exon 1, permet la transcription du gène dans les cellules hypophysaires lactotropes. Il exhibe entre autres des séquences de type « boîte E » permettant la liaison du facteur de transcription spécifique de l'hypophyse PIT1 (POU-homeodomain transcription factor 1) et est associé à des séquences de type enhancer et silencer.

On découvre ensuite un promoteur alternatif localisé 5,3 kb en amont. Ce promoteur distal, qui précède le TSS1 et l'exon alternatif 1a, est utilisé dans les sites de production extra-hypophysaires de la prolactine (cellules déciduales de l'endomètre utérin, cellules placentaires, cellules lymphoblastoïdes). Il exhibe des séquences CRE (cAMP-Responsive Element), une boîte CAAT de liaison aux protéines C/EBP et un motif Ets (AAGGAA) permettant la liaison des facteurs de transcription de la famille Ets. Ce promoteur est associé à une région enhancer en 5' capable de lier le facteurs AP-1 (JunD/c-Fos) via une séquence TRE (TPA DNA response element).

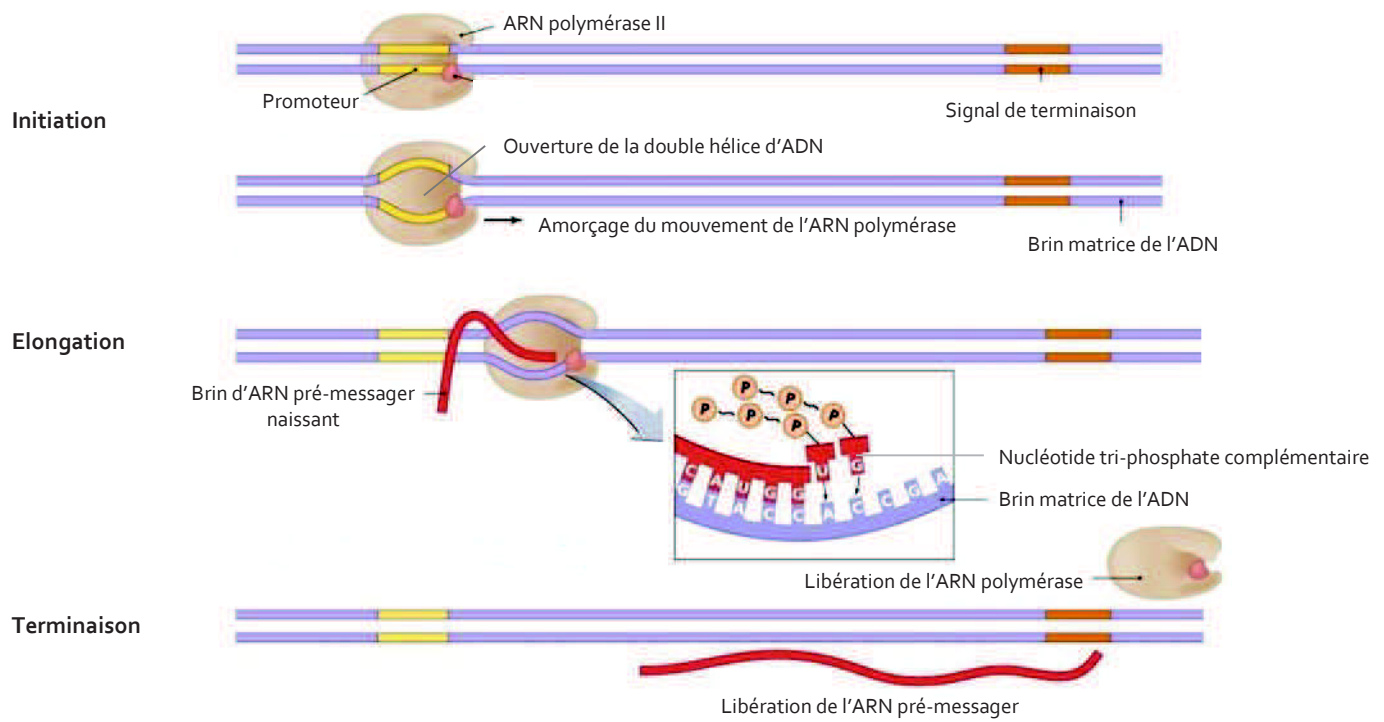
1.1.1.2. Unité de transcription

Un gène est formé d'une unité de transcription d'une part, et de séquences régulatrices d'autre part. L'unité de transcription est comprise entre un site d'initiation de la transcription (Transcription start signal : TSS) et des signaux de fin de transcription (signal de polyadénylation AATAAA notamment) ; c'est cette séquence qui sera transcrite, dans son intégralité, en ARN pré-messager. Au sein de l'unité de transcription, on distingue des séquences appelées exons et introns. Les séquences introniques transcrites, de tailles très variables, seront supprimées lors de la maturation de l'ARN messager et seules les séquences exoniques mises bout à bout lors de l'épissage constitueront l'ARN messager mature pris en charge par la machinerie traductionnelle. Au sein même des exons, l'ensemble de la séquence ne sera pas traduite par les ribosomes. C'est le cas des séquences situées en amont du codon « start » AUG de début de traduction (partie 5'UTR, pour UnTranslated Region) et en aval du codon « stop » de fin de traduction (partie 3'UTR). D'une longueur totale de 10,25 kb, la séquence génique de *PRL* a été étudiée dès la fin des années 1970 grâce au développement rapide des techniques de clonage (Cooke et al., 1981). L'unité de transcription du gène *PRL* (Figures 1C et D) a été initialement décrite comme composée de 5 exons et de 4 introns, grâce aux analyses de l'ARN messager hypophysaire (Truong et al., 1984). Néanmoins, on découvre au début des années 1990 l'existence d'un nouveau TSS et d'un exon non-codant, situés 5,3 kb en amont de l'exon 1 et utilisés dans les sites extra-hypophysaires de production de la prolactine (Gellersen et al., 1989).

1.1.1.3. Promoteur et séquences régulatrices

Le promoteur est une région d'ADN située en amont de l'unité de transcription d'un gène et qui permet d'initier sa transcription. Il comprend des séquences nucléotidiques spécifiques (les éléments cis géniques) qui permettront la liaison de facteurs trans protéiques, dont la machinerie transcriptionnelle et les facteurs de transcription. Le promoteur de cœur est défini comme la portion minimale d'un promoteur (environ 100 pb) nécessaire à l'initiation de la transcription. Celui-ci contient les sites de liaison de l'ARN polymérase et des facteurs généraux de transcription. La comparaison des promoteurs a permis de mettre en lumière l'existence de certaines séquences consensus, toutefois pas toujours présentes dans les promoteurs humains. Il s'agit par exemple de la boîte TATA (séquence riche en thymidines et adénosines) située vers -25 (25 pb en amont du TSS) et reconnue par le facteur général de transcription TATA-binding protein (TBP), ou encore de la séquence BRE (TFIIB Recognition Element) (séquence G/C G/C G/A C G C C) permettant la liaison de TFIIB (Smale and Kadonaga, 2003). Le promoteur de cœur est fréquemment prolongé en 5' par des séquences promotrices permettant la liaison de facteurs de transcription spécifiques (Levine and Tjian, 2003). Par exemple, la boîte CAAT vers -120 à -80 (séquence GGCCAATCT) est reconnue par les protéines C/EBP (CAAT enhancer-binding proteins) et la boîte E (séquence CACGTG) permet la liaison des facteurs de transcription à motif hélice-boucle-hélice (famille bHLH, pour basic Helix-Loop-Helix) tels que BMAL1-Clock, Pit-1 ou cMyc (Dong et al., 2011). Les gènes possèdent parfois des éléments régulateurs, amplificateurs (enhancers) ou inactivateurs (silencers), pouvant être situés à plusieurs kilobases du TSS, en amont ou en aval, voire sur un autre chromosome (Blackwood and Kadonaga, 1998 ; Fiering et al., 2000). Ces séquences régulatrices pourront également fixer des protéines capables de moduler l'activité du complexe transcriptionnel grâce au repliement de l'ADN dans le noyau (Westin and Schaffner, 1988). La multiplicité et l'hétérogénéité des séquences régulatrices encadrant un gène est un caractère primordial pour la régulation de l'expression du gène, par le jeu de l'expression et de la fixation différentielle de facteurs de transcription sur ces séquences.

A.



B.

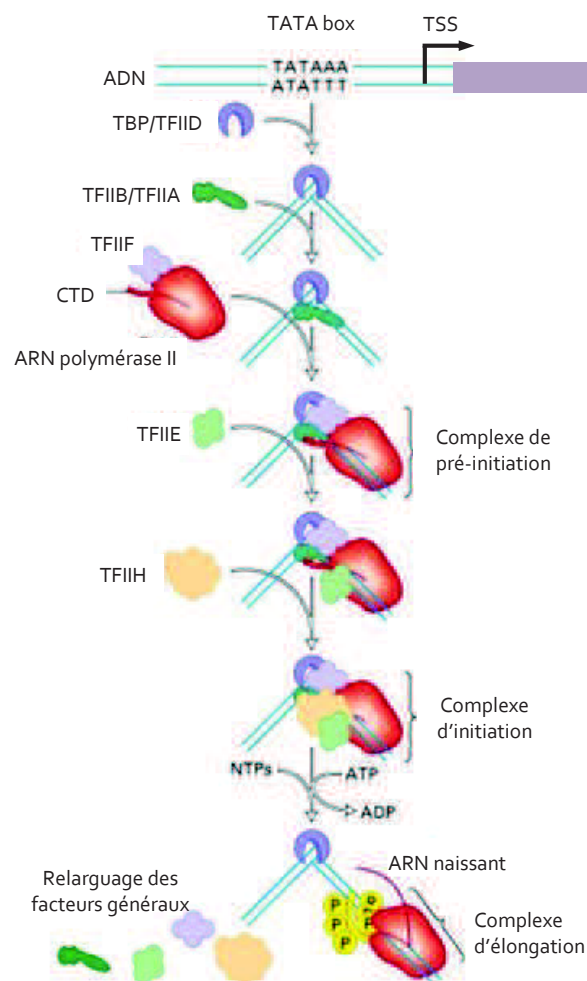


Figure 3 : Transcription d'un gène de classe 2

A. Les trois étapes de la transcription d'un gène. *Adapté de (Watson, 2006).*

B. Assemblage du complexe de pré-initiation de transcription. *Adapté de (Lodish, 2008).*

Le gène *PRL*, avec ses deux promoteurs alternatifs tissu-spécifiques, est un bon exemple de ces mécanismes de régulation (**Figure 2**) (Gerlo et al., 2006). On a tout d'abord décrit un premier promoteur, situé en amont du TSS₂ et de l'exon 1, qui permet la transcription du gène dans les cellules hypophysaires lactotropes. Il contient entre autres des séquences de type « boîte E » permettant la liaison du facteur de transcription spécifique de l'hypophyse PIT1 (POU-homeodomain transcription factor 1) et est associé à des séquences de type enhancer et silencer. On découvre ensuite un promoteur alternatif localisé 5,3 kb en amont (DiMattia et al., 1990). Ce promoteur distal, qui précède le TSS₁ et l'exon alternatif 1a, est utilisé dans les sites de production extra-hypophysaires de la prolactine (telles que les cellules déciduales de l'endomètre utérin ou les cellules lymphoblastoïdes) (DiMattia et al., 1998 ; Gellersen et al., 1994). Il présente des séquences CRE (cAMP-Responsive Element), une boîte CAAT de liaison aux protéines C/EBP et un motif Ets (AAGGAA) permettant la liaison des facteurs de transcription de la famille Ets (Brar et al., 2002). Ce promoteur est associé à une région enhancer en 5' capable de lier le facteur AP-1 (JunD/c-Fos) via une séquence TRE (TPA DNA Response Element) (Watanabe et al., 2001). L'utilisation de promoteurs alternatifs n'est pas un phénomène rare, puisqu'on estime qu'entre 18% (Landry et al., 2003) et 50% (Kimura et al., 2006) des gènes humains seraient transcrits via des promoteurs multiples.

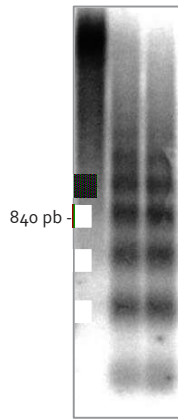
1.1.2. La transcription, étape fondamentale de l'expression d'un gène

La transcription d'une séquence génique en ARN pré-messager est la première et la plus fondamentale des étapes de l'expression d'un gène. Dans les études transcriptomiques, c'est la présence de l'ARN messager qui va signer « l'expression » d'un gène donné. Néanmoins, il ne faut pas négliger les étapes ultérieures (maturation de l'ARN messager, traduction, modifications post-traductionnelles...) qui pourront également conditionner l'abondance ou la stabilité du produit final.

La transcription des gènes de classe 2 est réalisée par un complexe de transcription incluant l'ARN polymérase de type II et des facteurs généraux de la transcription (Thomas and Chiang, 2006) qui permettent le bon déroulement des différentes étapes de la transcription (initiation, élongation, terminaison) (**Figure 3A**). La première étape de la transcription, l'initiation, est particulièrement cruciale car elle va directement conditionner le taux de transcription d'un gène (**Figure 3B**). Dans un premier temps, la protéine TBP (sous-unité du complexe TFIID) se fixe à la séquence nucléotidique de la boîte TATA (Tjian, 1996). Puis cinq autres facteurs généraux d'initiation de la transcription ainsi que l'ARN polymérase s'associent pour former le complexe de pré-initiation (CPI). La phosphorylation de la queue C-terminale de l'ARN polymérase (CTD pour carboxy-terminal domain) par le facteur TFIIH à activité kinase semble impliquée dans la transition vers l'étape d'élongation (Hirose and Ohkuma, 2007). La machinerie transcriptionnelle (ARN polymérase associée à des facteurs généraux de type « élongateurs ») parcourt alors le brin matrice de l'ADN de l'extrémité 3' vers l'extrémité 5', en ajoutant à l'ARN pré-messager naissant des nucléotides complémentaires aux nucléotides du brin matrice (Palangat and Larson, 2012). Les mécanismes de terminaison de la transcription ne sont pas parfaitement connus mais impliquent la destruction des liaisons hydrogènes entre ADN et ARN libérant le brin d'ARN synthétisé, ainsi que le clivage et l'ajout de nombreuses adénosines (polyadénylation) à l'extrémité 3' de l'ARN messager (Kuehner et al., 2011). Seul un niveau basal de transcription sera obtenu avec le complexe de pré-initiation seul. Ce sont d'autres protéines de type activateur et/ou répresseur qui vont être capables de moduler ce taux de transcription basal en se liant aux éléments régulateurs (**voir paragraphe 1.2**).

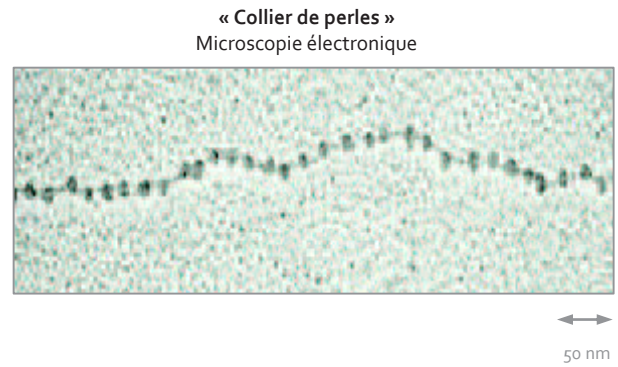
La transcription du gène *PRL* peut aboutir à deux variants ARN distincts selon le promoteur et le TSS qui ont été utilisés par la machinerie transcriptionnelle (**Figure 1D**). Le transcrit *PRL* placentaire (variant 2) est formé de 6 exons, incluant l'exon 1a non-codant absent du transcrit hypophysaire (variant 1) (**Figure 1E**) (Gellersen et al., 1989). Entre ces deux transcrits, seule la partie 5'UTR non codante diffère : la séquence codante est identique et la même protéine sera synthétisée lors de la traduction (**Figure 1F**).

A.

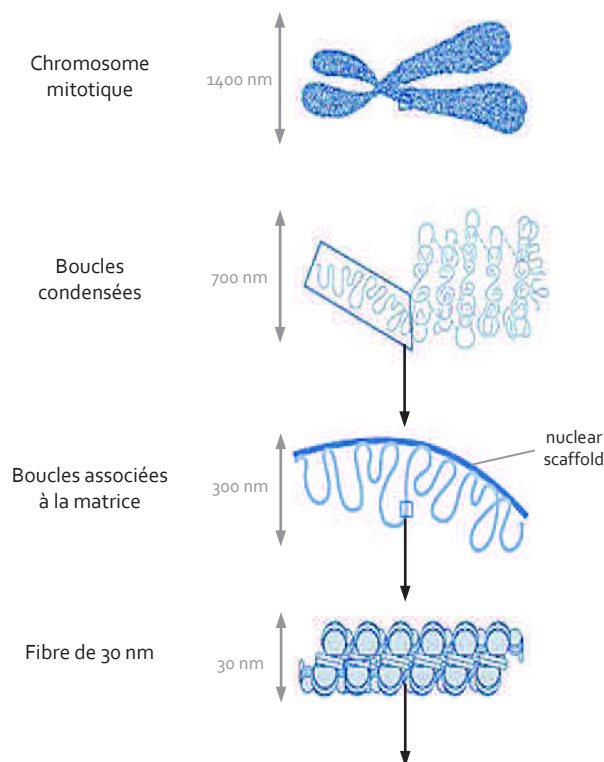


Migration électrophorétique de fragments d'ADN après digestion de chromatine par nucléase

B.



C.



D.

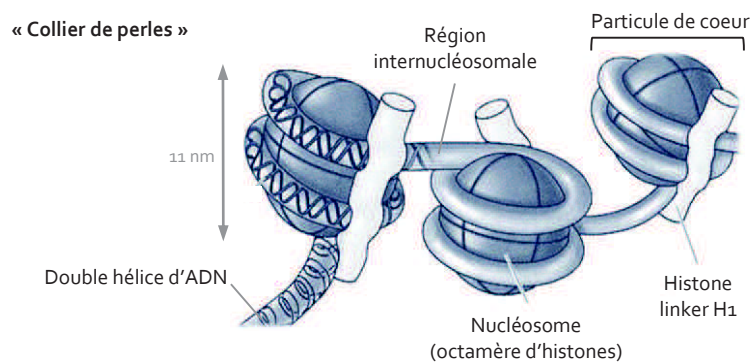


Figure 4 : La structure chromatinienne

A. Observation des fragments d'ADN « en échelle » après digestion partielle de la chromatine. *Tiré de (Koop et al., 2003).*

B. Observation d'une fibre de chromatine en microscopie électronique.

C. Les états de compaction de la chromatine. *Inspiré de la figure « Hierarchical folding of chromatin » (Lodish, 2008).*

D. L'ADN nucléosomale.

1.1.3. Le paysage épigénétique contrôle l'accessibilité de la chromatine

1.1.3.1. Le nucléosome, unité structurale fondamentale de la chromatine

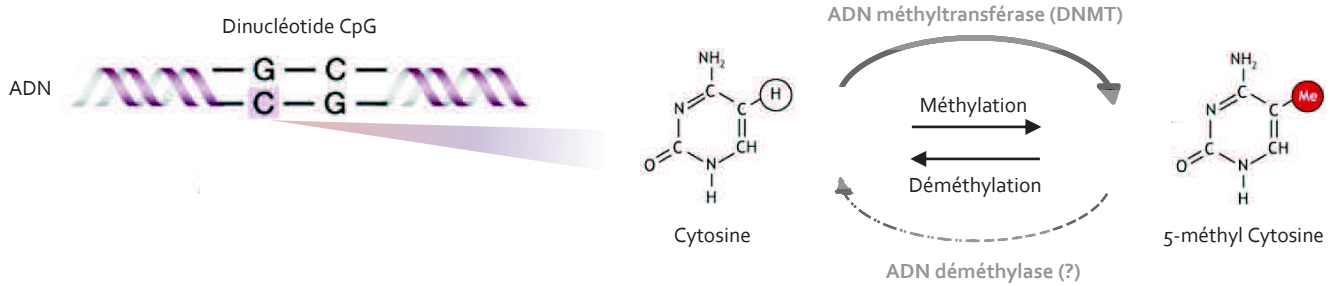
Dans le noyau de la cellule, la molécule d'ADN n'existe pas sous la forme d'une molécule libre. Le génome nucléaire s'organise en une structure nucléoprotéique, appelée chromatine, association de la double hélice d'ADN et des protéines associées à l'ADN. C'est donc la chromatine, et non l'ADN seul, qui est impliquée dans tous les événements moléculaires faisant intervenir le matériel génétique, tels que la réplication, la réparation, la transcription ou la ségrégation de l'ADN lors des divisions cellulaires. Historiquement, la nature de la chromatine a été identifiée par des méthodes biochimiques et par microscopie électronique. Dans les années 1970, on observe que la digestion partielle de la chromatine génère des fragments d'ADN résiduels d'environ 200 paires de bases et leurs multiples (Hewish and Burgoyne, 1973 ; Williamson, 1970) (Figure 4A). Les progrès de la microscopie électronique permettent ensuite l'observation de la chromatine et révèlent des particules espacées à intervalle régulier, en "collier de perles" (Olins and Olins, 1974 ; Oudet et al., 1975) (Figure 4B): ces particules sont nommées nucléosomes. En parallèle, des analyses biochimiques permettent d'isoler un complexe de 200 kDa, composé d'ADN et de protéines histones associées à l'ADN (Kornberg and Thomas, 1974). Finalement, la structure du nucléosome est déterminée par cristallographie aux rayons X (Finch et al., 1977 ; Luger et al., 1997). Le nucléosome se compose d'une particule de cœur et d'une région de liaison internucléosomale (Figure 4D) (Richmond and Davey, 2003). La région de liaison séparant deux nucléosomes est caractérisée par la présence de l'histone H1. La particule de cœur se compose d'un octamère protéique autour duquel s'enroule la double hélice d'ADN (146 paires de bases d'ADN enroulé 1,7 fois autour de l'octamère). Cet octamère contient 2 copies de chacune des histones H3, H4, H2A et H2B. Les histones de cœur sont des protéines de petite taille (12-16 kDa, 100 à 130 acides aminés) et basiques (composées à plus de 20 % des résidus arginine et lysine, elles possèdent une charge globale positive). La présence de ces charges positives leur permet d'interagir avec le squelette phosphodiester de l'ADN chargé négativement. La région la plus conservée est la partie centrale, comportant le domaine "histone fold" qui permet aux histones d'interagir entre elles (en dimères H2A-H2B et H3-H4) via des interactions de nature hydrophobe de type "poignée de mains" (Arents et al., 1991). Contrairement au domaine central, les queues amino-terminales des histones de cœur sont très variables et peu structurées ; elles émergent du cylindre compact formé par les domaines globulaires et sont le site de nombreuses modifications post-traductionnelles (Khorasanizadeh, 2004) (voir paragraphe 1.1.3.3.2).

L'association entre double hélice d'ADN et octamère d'histones constitue le premier niveau d'organisation de la chromatine, l'ADN nucléosomal. Cependant, il existe des niveaux supérieurs de structuration de la chromatine, plus mal connus (Figure 4C), tels que la fibre de 30 nm (enroulement du "collier de perles" en une structure cylindrique d'un diamètre de 30 nm) (Finch and Klug, 1976), elle-même repliée en boucles le long d'une matrice protéique centrale appelée matrice nucléaire ("nuclear scaffold"). L'association de tels ensembles forme des sous-domaines chromosomiques (Cremer and Cremer, 2001), dont la somme aboutit au chromosome mitotique qui représente le niveau de compaction maximal de la chromatine, permettant la bonne ségrégation du matériel génétique entre deux cellules-filles lors de la division cellulaire.

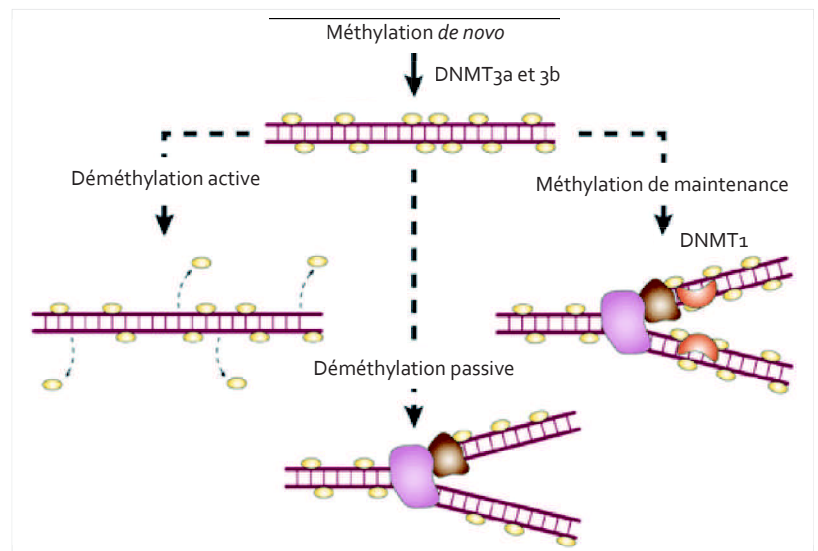
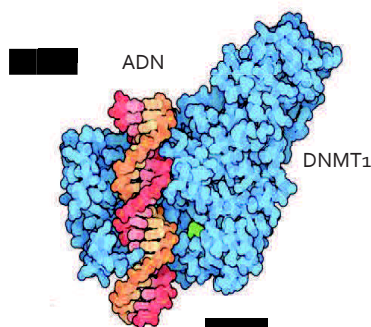
1.1.3.2. Dynamique de l'organisation chromatinienne

Historiquement, sur la base des observations au microscope électronique du noyau des cellules, la chromatine fut divisée en deux domaines distincts : l'hétérochromatine localisée en périphérie du noyau et dense aux électrons, et l'euchromatine « structure claire et décondensée au cours de l'interphase » (Heitz, 1928). L'hétérochromatine elle-même peut être divisée en deux sous-catégories, l'hétérochromatine constitutive fortement compactée qui contient peu de gènes et est formée principalement de séquences répétées situées à proximité des centromères et des télomères, et l'hétérochromatine facultative qui contient des régions codantes qui, bien qu'elles adoptent les caractéristiques structurales et fonctionnelles de

A. Modification



B. Enzymes « écrivains » et « effaceurs »



C. Protéines « lecteurs »

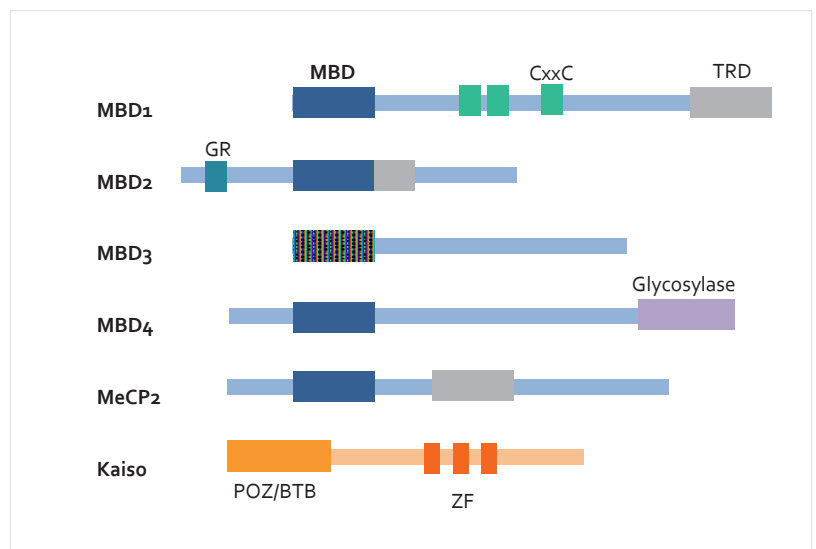
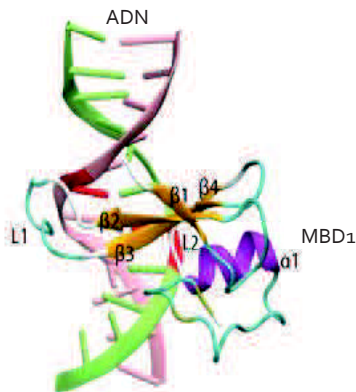


Figure 5 : Méthylation de l'ADN

A. La cytosine d'un dinucléotide CpG de l'ADN peut être modifiée par ajout d'un groupement méthyle sur son carbone 5.

B. Les DNMT s'associent à l'ADN et catalysent la méthylation des cytosines.

(Panneau de gauche) Structure de la protéine DNMT1 associée à la double hélice d'ADN. Tiré de (Goodsell, 2011).

(Panneau de droite) Les DNMT3 réalisent la méthylation *de novo* et DNMT1 assure la méthylation de maintenance par méthylation du brin néosynthétisé lors de la réplication. La déméthylation peut être active, catalysée par une ADN déméthylase encore discutée, ou passive par dilution lors de la réplication. Traduit de (Reik and Walter, 2001).

C. Les MBP (Methyl Binding Proteins) sont capables de lier l'ADN méthylé via leur MBD (Methyl Binding Domain).

(Panneau de gauche) Structure du domaine MBD1 associé à une double hélice d'ADN. Tiré de (Zou et al., 2012).

(Panneau de droite) TRD : Domaine de répression de la transcription. CxxC : motif de liaison à l'ADN CxxC. GR : région riche en résidus glycine et arginine. Glycosylase : domaine thymine glycosylase. ZF : motif de liaison à l'ADN à doigt de zinc.

l'hétérochromatine, peuvent se décompacter sous certaines conditions. L'état de compaction a été corrélé à l'activité transcriptionnelle au sein de ces deux domaines. La transcription est majoritairement réprimée au sein de l'hétérochromatine, alors que l'euchromatine a été décrite comme associée à une transcription active des gènes.

En effet, dès le premier niveau de compaction (ADN nucléosomal), l'information génétique n'est déjà plus totalement accessible. L'enroulement de l'ADN autour de l'octamère d'histones constitue une barrière physique naturelle pour les machineries nucléaires nécessitant d'avoir accès au matériel génétique (Fyodorov and Kadonaga, 2001). L'accessibilité physique de la séquence génique à la machinerie transcriptionnelle va donc être nécessaire à l'expression d'un gène (Carey, 2000). La cellule possède des mécanismes lui permettant de modifier localement la structure et le degré de compaction de la chromatine. Ceci illustre le fait que la chromatine n'est pas une structure homogène et figée dans le temps, mais une structure dynamique qui peut se présenter sous différentes conformations permettant ou non le déroulement d'activités nucléaires comme la transcription. Cette composante structurale, dynamique et régulée, qui s'ajoute à l'information génétique portée par l'ADN, est dirigée par des processus dits « épigénétiques » qui vont affecter l'environnement du génome sans modifier la séquence en nucléotides de l'ADN. La composante épigénétique va donc décrire la manière dont les modifications de l'environnement chromatinien vont contrôler le niveau de compaction de la chromatine et ainsi affecter les expressions géniques.

1.1.3.3. Marques et acteurs épigénétiques

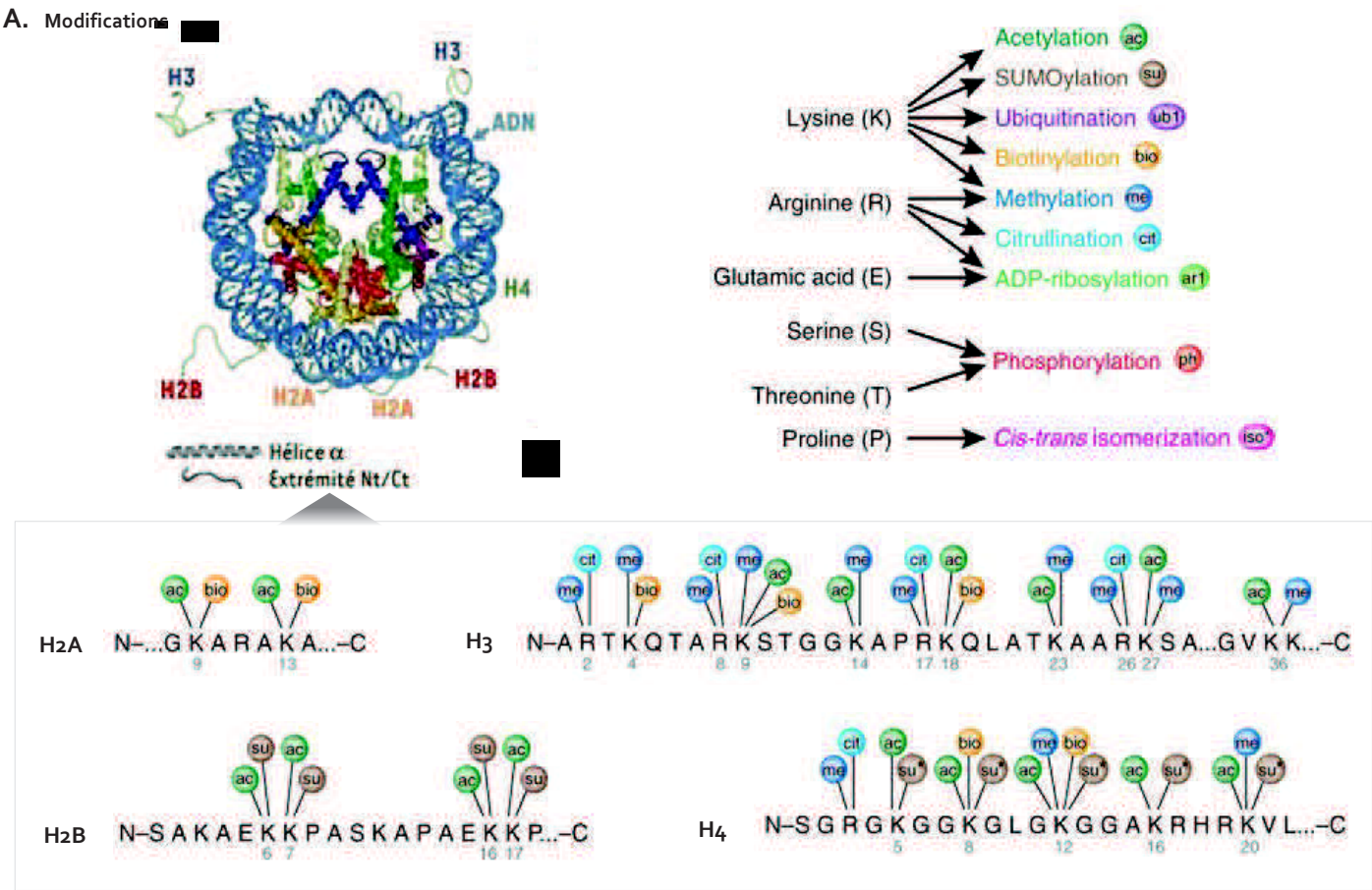
Les nucléosomes, unités de base du matériel génétique sont, au niveau de leur composition comme de leur arrangement, la cible principale des régulations épigénétiques. Les marques épigénétiques permettent de moduler l'emplacement des nucléosomes, le degré de compaction local de la chromatine, le recrutement d'effecteurs épigénétiques, et finalement l'activité transcriptionnelle des gènes. L'information épigénétique au sein de la chromatine est principalement véhiculée par des modifications de l'ADN et des histones (Sawan et al., 2008) – ce sont les deux types de marques que vous trouverez décrites ci-après – en association notamment avec l'incorporation de variants d'histones ou la régulation par les miARN.

1.1.3.3.1. La méthylation de l'ADN

La méthylation de l'ADN correspond à l'ajout d'un groupement méthyle ($-CH_3$) fixé de manière covalente sur une base azotée de l'ADN (Figure 5A). Chez les eucaryotes, la méthylation a lieu sur le carbone 5 d'un résidu cytosine (C) de l'ADN localisé en 5' d'un résidu guanine (G) dans un dinucléotide CpG. Les dinucléotides CpG sont répartis de façon non uniforme dans le génome, avec une sous-représentation générale, mais un enrichissement dans de courtes régions, appelées « îlots CpG » (Bird, 2002). Ces « îlots CpG », zones de l'ADN avec un fort contenu en nucléotides G et C ($GC > 50\%$) et une forte fréquence des dinucléotides CpG ($CpG > 0.6$), sont notamment localisés au niveau des promoteurs des gènes (Gardiner-Garden and Frommer, 1987).

La méthylation d'une cytosine est assurée par une famille d'enzymes appelées ADN méthyltransférases (DNMT) (Figure 5B). On connaît 5 DNMT classées selon leur homologie de séquence dans leur partie catalytique carboxy-terminale. DNMT1 est impliquée dans la méthylation de maintenance au fil des divisions successives, par méthylation du brin néo-synthétisé lors de la réplication (Rountree, 2000). DNMT3a et DNMT3b sont les méthyltransférases impliquées dans la méthylation *de novo* lors du développement des cellules germinales et lors de l'embryogenèse précoce (Goll and Bestor, 2005). Le processus inverse de déméthylation de l'ADN est un phénomène débattu. En effet, il n'est pas encore clair si la déméthylation est un phénomène passif de dilution au cours des divisions cellulaires, ou si la réversibilité de la réaction de méthylation est sous la dépendance d'une déméthylase. La protéine MBD2 a été proposée comme déméthylase de l'ADN (Bhattacharya et al., 1999 ; Detich et al., 2002), mais ces résultats sont discutés (Ng et al., 1999). La méthylation de la cytosine est considérée chez l'Homme comme une marque associée à l'hétérochromatine et

A. Modifications



B. Enzymes « écrivains » et « effaceurs »

Modification	Résidu cible	Effecteur	
		qui ajoute la modification	qui enlève la modification
Acetylation	Lysines	Histone acétyltransférase (HAT)	Histone déacétylase (HDAC)
Methylation	Lysines Arginines	Histone méthyltransférase (HMT)	Histone déméthylase (HDM)
Phosphorylation	Sérines	Kinase	Phosphatase
Ubiquitination	Lysines	Ubiquitine ligase (E1, E2, E3)	Isopeptidase

C. Protéines « lecteurs »



Figure 6 : Modifications post-traductionnelles des histones

A. Ce sont principalement les résidus des queues N-terminales des histones qui sont la cible de modifications post-traductionnelles. De nombreux types de groupements peuvent être ajoutés. *Adapté de (Latham and Dent, 2007).*

B. Différentes familles d'enzymes assurent l'ajout et le retrait des marques d'histones.

C. Domaines protéiques (et protéines) reconnaissant spécifiquement certains résidus modifiés des queues d'histones. *Inspiré de (Kouzarides, 2007b).*

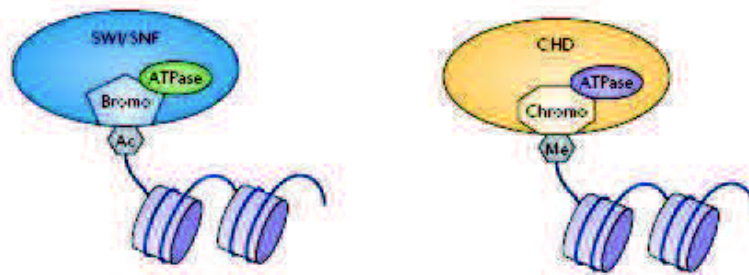
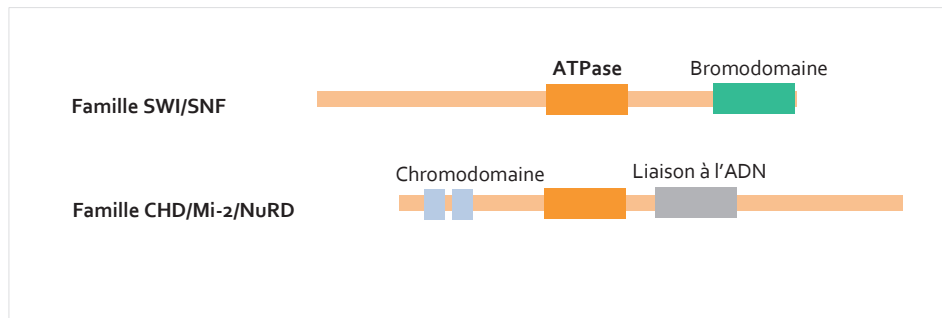
cette marque épigénétique est classiquement responsable d'une chromatine transcriptionnellement silencieuse (**Figure 8B**) (Bird and Wolffe, 1999 ; Wade, 2001). Ainsi, chez l'Humain, 76% des promoteurs des gènes contiennent des îlots CpG (Bird, 2002) et leur méthylation est généralement une marque forte de répression de la transcription (Davuluri et al., 2001).

1.1.3.3.2. Les modifications des histones

Au-delà des marques apposées directement sur l'ADN (méthylation des cytosines), les protéines histones constituent également des cibles privilégiées pour les modifications épigénétiques. Elles peuvent subir des modifications post-traductionnelles, covalentes et réversibles, pour la plupart au niveau de l'extrémité N-terminale non structurée qui émerge à la surface du nucléosome. Différentes modifications ont été décrites comme l'acétylation, la phosphorylation, la méthylation, l'ubiquitylation, la sumoylation (Kouzarides, 2007b) (**Figure 6A**). Plus récemment, le champ d'investigation concernant les modifications post-traductionnelles d'histones s'est considérablement élargi suite à la découverte de nombreuses nouvelles marques (Tan et al., 2011). Ces modifications peuvent affecter une soixantaine d'acides aminés des quatre histones canoniques, dont des résidus lysine, arginine, glutamate, sérine et thréonine (Kouzarides, 2007b). Un même acide aminé peut subir un très grand nombre de modifications distinctes. La lysine 12 de l'histone H4 peut par exemple être détectée sous six états différents (non modifiée, acétylée, biotinylée, mono- di- ou tri- méthylée). Le nombre de combinaisons théoriques possibles au sein d'un seul nucléosome est évalué à plus d'un milliard de milliards de combinaisons. Les modifications post-traductionnelles des histones fonctionnent de manière combinatoire, comme un code (Jenuwein and Allis, 2001). Ce « code histone » permettrait d'associer à un jeu de modifications post-traductionnelles un état particulier de la chromatine, d'en réguler la compaction et l'accessibilité, et d'en modifier les fonctions essentielles. Certaines modifications post-traductionnelles des histones ont été décrites comme caractéristiques de l'hétérochromatine, d'autres de l'euchromatine (**Figure 8A et B**). L'état hyperacétylé des histones correspond généralement aux régions moins compactées et actives du génome (Felsenfeld et al., 1996). La triméthylation H3K4 (résidu lysine 4 de l'histone H3) est également considérée comme la marque d'une chromatine permissive à la transcription (Noma et al., 2001). *A contrario*, l'hypoacétylation et la méthylation des histones (dont les di- et tri-méthylation d'H3K9 ou la tri-méthylation d'H3K27) sont des marques de la chromatine plus compactée n'autorisant pas la transcription des séquences géniques voisines.

L'ajout ou la suppression de ces modifications sont catalysées par des enzymes « écrivains » et « effaceurs », souvent spécifiques d'une modification et d'un résidu donné, qui sont peu à peu décrites (**Figure 6B**). Le niveau d'acétylation des histones est ainsi la résultante de l'activité de deux familles d'enzymes antagonistes : les histone-acétyltransférases (HAT) qui ajoutent un groupement acétyle sur une lysine à partir de l'acétyl-coenzyme A et les histone-désacétylases (HDAC) qui enlèvent ce groupement acétyle. Les HAT peuvent être divisées en trois familles (GNAT, MYST et CBP/P300) (Sternier and Berger, 2000) alors que 18 HDAC humaines ont été identifiées et réparties en 3 classes (I, II et III) (Thiagalingam et al., 2003). Les histone-méthyltransférases (HMT) catalysent quant à elles la méthylation des lysines et des arginines : c'est le cas des protéines à domaine SET (Suppressor of the Variegation - Enhancer of Zeste and Trithorax) (Morgunkova and Barlev, 2006). Quelques histone déméthylases (HDM) ont été identifiées à ce jour, dont la protéine LSD1 (Lysine Specific Demethylase 1) d'une part (Shi et al., 2004), et les protéines contenant un domaine JmJC (Jumonji C) d'autre part (Whetstine et al., 2006 ; Yamane et al., 2006). La réversibilité des modifications d'histone, tout comme celle de la méthylation des CpG de l'ADN, met en exergue la plasticité de l'information épigénétique, face à la stabilité du code génétique.

A.



B.

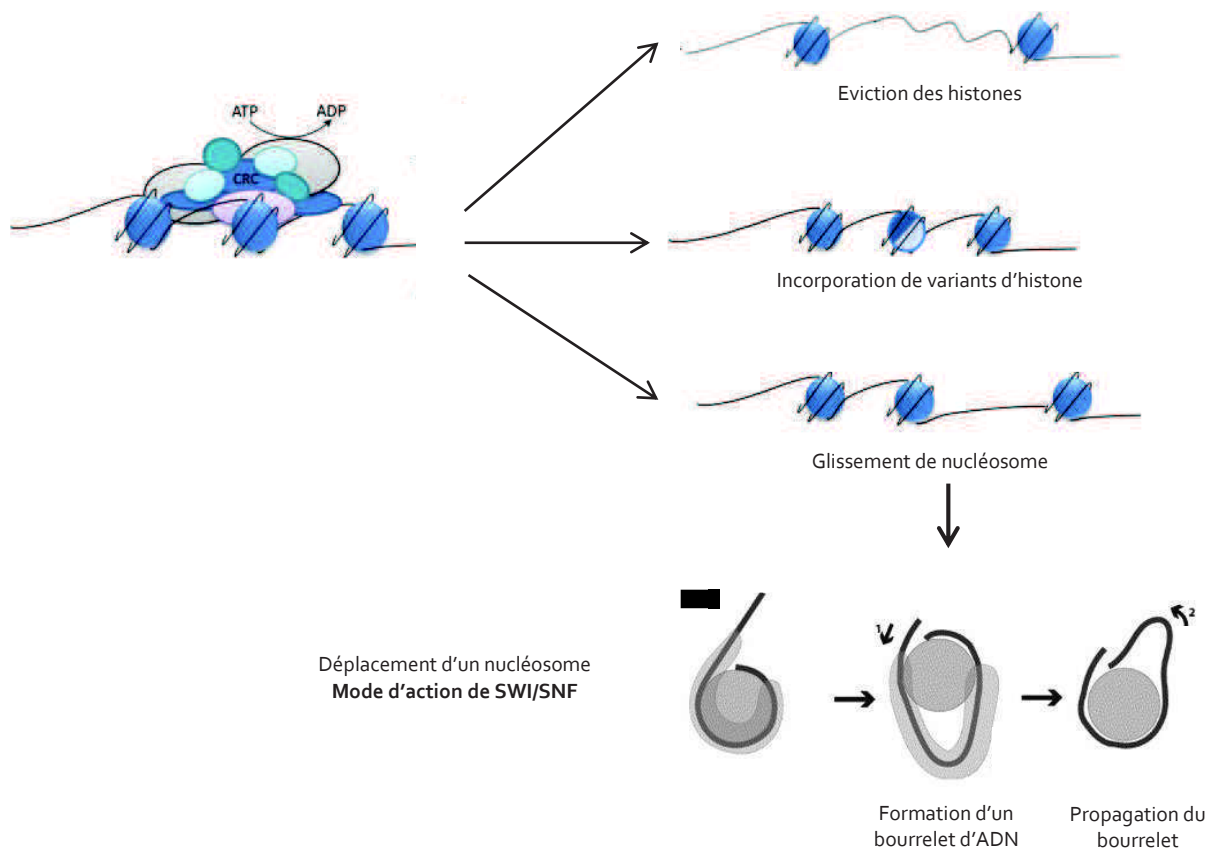


Figure 7 : Complexes de remodelage de la chromatine

A. Chaque famille contient un domaine ATPase et un motif de reconnaissance à la chromatine. Tiré de (de la Serna et al., 2006).

B. Les complexes de remodelage de la chromatine (CRC), assemblages multi-protéiques se liant à l'ADN, sont capables de remanier les nucléosomes en utilisant l'énergie d'hydrolyse de l'ATP. Traduit de (Gangaraju and Bartholomew, 2007).

1.1.3.3. Les complexes de remodelage de la chromatine

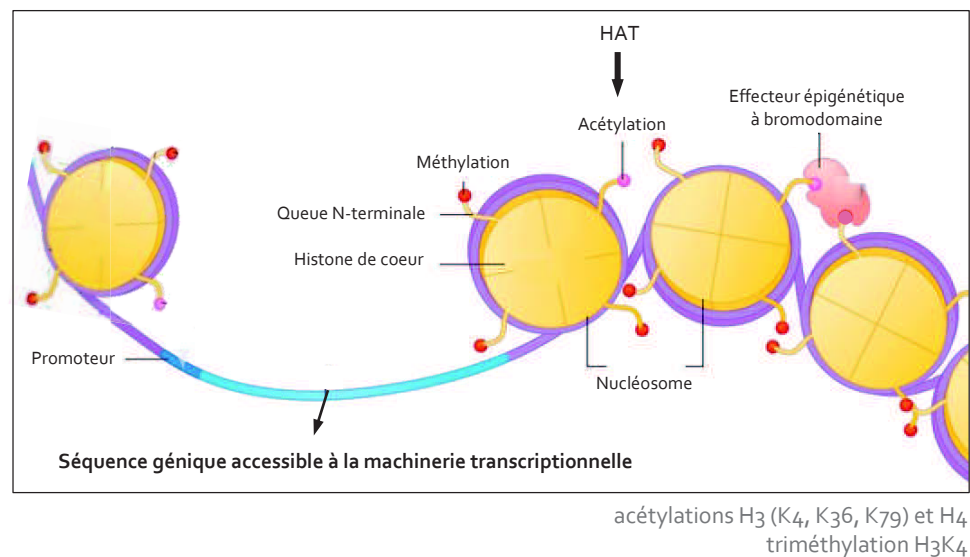
L'initiation de la transcription d'un gène se déroule dans un contexte très restrictif, car la structure nucléosomale est physiquement peu propice à la liaison des complexes protéiques de haut poids moléculaire formant le complexe de pré-initiation. Toutefois, comme aucun contact direct n'existe entre les bases azotées de l'ADN et les histones, le nucléosome ne présente aucune spécificité à la séquence d'ADN et sa position sur le génome peut varier. La cellule possède des complexes de remodelage de la chromatine, capables de mobiliser ces nucléosomes (**Figure 7**). Il existe cinq familles de facteurs de remodelage qui possèdent tous une sous-unité ATPase qui constitue le moteur du complexe car le déplacement des nucléosomes utilise l'énergie fournie par l'hydrolyse de l'ATP (**Figure 7A**). Les facteurs de la famille SWI/SNF (mating type Switching/ Sucrose Non Fermenting) espacent les nucléosomes au niveau des promoteurs des gènes ([Boeger et al., 2003](#) ; [Martens and Winston, 2003](#) ; [Reinke and Horz, 2003](#)) rendant les séquences d'ADN régulatrices accessibles aux facteurs d'initiation de la transcription (**Figure 7B**). Le mode d'action particulier de ce complexe reposerait sur le déplacement d'un nucléosome obtenu en poussant un surplus d'ADN à l'entrée du nucléosome puis en le faisant glisser ([Narlikar et al., 2002](#) ; [Saha et al., 2005](#) ; [Zofall et al., 2006](#)). Ce complexe de remodelage peut être recruté sur les résidus d'histones acétylés via son bromodomaine ([Chatterjee et al., 2011](#) ; [Hassan et al., 2001](#)) (**Figure 7A**). A l'inverse, le complexe NuRD (Nucleosome Remodeling and Deacetylation) est associé à une forte répression transcriptionnelle ([Denslow and Wade, 2007](#)). Son mode d'action repose d'une part sur le remodelage médié par les sous-unités Chd3 et Chd4 qui induisent le glissement ATP-dépendant des nucléosomes le long de l'ADN ([Brehm et al., 2000](#) ; [Wang and Zhang, 2001](#)) et d'autre part sur l'activité des HDAC intégrées à ce complexe capables de déacétyler localement les histones. Les complexes de remodelage de la famille NuRD peuvent être ciblés sur les cytosines méthylées de l'ADN via un domaine MBD (Methyl Binding Domain) ([Bowen et al., 2004](#)), mais également sur les histones méthylées via un chromodomaine ou un domaine PHD (Plant HomeoDomain).

1.2. Mécanismes de régulation des expressions géniques : comment activer ou réprimer un gène ?

La régulation de l'expression d'un gène fait référence au contrôle qualitatif (quand ?) et quantitatif (combien ?) de l'apparition du produit fonctionnel d'un gène. Pour les gènes humains de classe 2 codant des protéines, il existe différents niveaux de contrôle : modulation de l'activité transcriptionnelle, régulation post-transcriptionnelle (contrôle de l'épissage, de l'export nucléaire ou de la stabilité de l'ARN messenger), ARN interférence, contrôle traductionnel et régulations post-traductionnelles (contrôle du repliement, des modifications chimiques ou de la stabilité de la protéine). Tous ces niveaux peuvent se combiner pour finalement arriver à un contrôle fin de la production d'une protéine en terme de temps, de quantité, de localisation ou d'activité. La régulation transcriptionnelle est le premier niveau de contrôle. Bien que possédant l'ensemble de l'information génétique de l'organisme, une cellule humaine différenciée ne transcrit pas tous les gènes contenus par son génome. Nous nous attacherons ici à décrire comment une cellule est capable d'activer certains gènes et d'en réprimer d'autres, plus particulièrement en contrôlant le niveau de production de l'ARN pré-messager lors de la transcription. Cette régulation transcriptionnelle, qu'elle soit activatrice ou répressive, utilise deux leviers principaux : un levier génétique impliquant une interaction entre une séquence génique régulatrice et une protéine modulant l'activité de la machinerie transcriptionnelle ; et un levier épigénétique qui implique l'apposition ou l'effacement de marques épigénétiques, le recrutement de protéines effectrices sur ces marques ou l'action de complexes de remodelage conditionnant localement la structure chromatiniennne.

A.

Chromatine en conformation « ouverte »



B.

Chromatine en conformation « fermée »

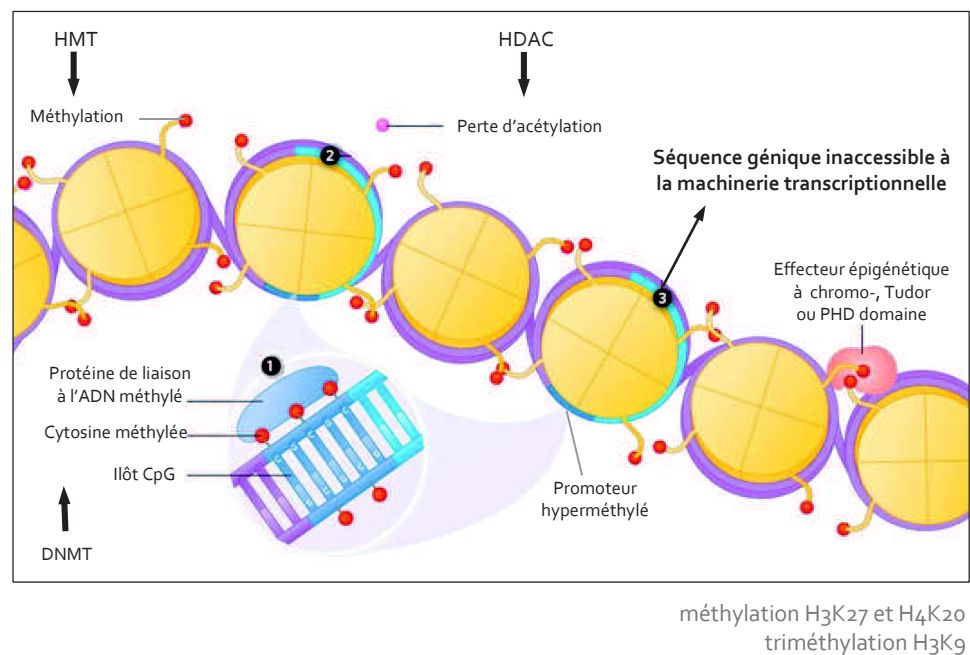


Figure 8 : Marques épigénétiques, états chromatinien et permissivité à la transcription

A. Les marques associées à l'activation transcriptionnelle sont la méthylation de H3K4, H3K36 et H3K79 ; K4 et K36 étant impliquées dans l'élongation (Kouzarides, 2007a).

B. Les marques associées à la répression de la transcription sont quant à elles la méthylation de H3K9 (bien que sa forme triméthylée soit présente au sein de la région codante des gènes transcriptionnellement actifs), H3K27 et H4K20 (Kouzarides, 2007a).

1.2.1. Mécanismes d'activation transcriptionnelle

Les divers mécanismes d'activation transcriptionnelle tendent à augmenter la production de l'ARN pré-messager (**Figure 9**). On distinguera dans ce paragraphe les mécanismes nés de l'interaction spécifique d'une protéine (facteur de transcription) sur une séquence cis de l'ADN d'une part et les mécanismes issus de la présence de marques épigénétiques (en particulier l'acétylation des histones) d'autre part.

1.2.1.1. Facteurs de transcription spécifiques à activité activatrice

Les facteurs de transcription spécifiques sont des protéines capables d'interagir avec l'ADN de façon séquence-spécifique au niveau d'une région régulatrice ([Carey, 2000](#)). Ce ciblage à une séquence ADN particulière est assuré par des domaines de liaison à l'ADN tels que le motif Hélice-boucle-hélice (présent chez Pit-1, Myc ou la famille Ets), le domaine en doigt de zinc (chez Snail ou KRAB-ZFP) ou le domaine « glissière à leucine » (Leucine zipper) (chez AP-1 ou C/EBP). Certains facteurs de transcription ont besoin d'une modification post-traductionnelle pour être actifs. L'activité de Ets-1 est par exemple régulée par de multiples modifications : la phosphorylation par CAMKII empêche sa fixation à l'ADN alors que la phosphorylation médiée par la MAPK permet à Ets-1 de recruter le coactivateur CBP/p300 et d'activer la transcription ([Tootle and Rebay, 2005](#)). Quant au facteur de transcription p53, son affinité pour l'ADN augmente lorsqu'il est acétylé ([Gu and Roeder, 1997](#)). D'autres facteurs, incluant les récepteurs nucléaires des glucocorticoïdes, doivent être complexés avec leur ligand pour assurer leur fonction de transactivation ([Wessely et al., 1997](#)). Certains facteurs de transcription sont spécifiques d'un type cellulaire particulier. C'est le cas de Pit-1, présent uniquement dans les cellules de l'antéhypophyse, responsable de la production des hormones hypophysaires. Il se fixe sur le promoteur proximal du gène *PRL* via son domaine hélice-boucle-hélice et en active la transcription ([Cohen et al., 1996](#)).

Une fois liés à l'ADN, les facteurs de transcription à activité activatrice induisent ou facilitent la transcription des gènes ([Carey, 2000](#)). La capacité transactivatrice d'un facteur de transcription peut tout d'abord reposer sur une interaction avec la machinerie transcriptionnelle. Certains facteurs de transcription vont ainsi recruter et/ou stabiliser la liaison du complexe d'initiation de la transcription au promoteur ([Beato and Eisefeld, 1997](#)) par exemple en accélérant l'assemblage des facteurs de transcription généraux au niveau du promoteur. D'autres vont favoriser la transition entre le complexe de pré-initiation et le complexe compétent pour l'élongation. Cette transactivation peut être assurée par une interaction directe, via des coactivateurs ou par l'intermédiaire d'un complexe multi-protéique. Ainsi le complexe TRAP/DRIP/Mediator, lié aux récepteurs nucléaires ([Rachez and Freedman, 2000](#)), active la transcription via une interaction directe avec l'ARN polymérase II, notamment en phosphorylant son domaine C-terminal (**Figure 9**).

La capacité transactivatrice d'un facteur de transcription peut également reposer sur l'ouverture locale de la chromatine, rendant la séquence promotrice accessible à la machinerie transcriptionnelle et à d'autres facteurs de transcription ([Eberharter and Becker, 2002 ; Grunstein, 1997](#)). Cette plasticité de la chromatine est permise par le recrutement par le facteur de transcription d'acteurs épigénétiques, tels que les complexes de remodelage de type SWI/SNF qui altèrent l'organisation spatiale des nucléosomes au niveau du promoteur, ou des protéines qui vont ajouter des modifications chimiques aux queues N-terminales des histones (**Figure 9**). Ainsi, la HAT CBP/p300 (CREB Binding Protein/p300) agit fréquemment comme coactivateur transcriptionnel. Ses différents domaines de liaison lui permettent d'interagir avec de nombreux facteurs de transcription (CREB (cAMP Response Element Binding protein), AP-1, p53, Ets-1...) et d'acétyler les histones à proximité ([Arias et al., 1994 ; Bannister and Kouzarides, 1996 ; Ogryzko et al., 1996](#)).

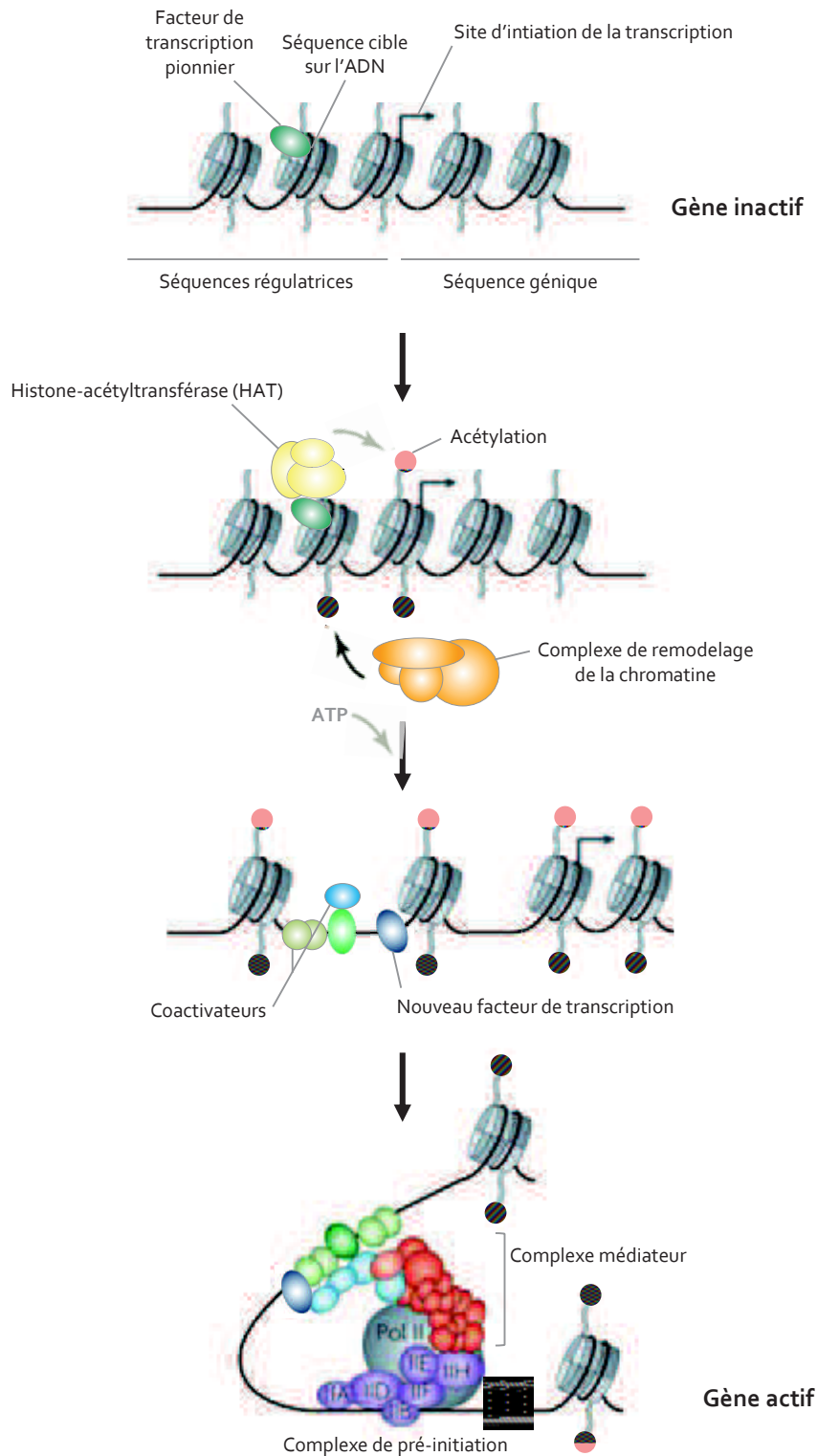


Figure 9 : Mécanismes d'activation transcriptionnelle

Inspiré de l'activation transcriptionnelle du gène HO chez la levure (Bhoite et al., 2001 ; Malik and Roeder, 2010).

Il s'agit ici de la représentation des différentes étapes d'un modèle générique d'activation transcriptionnelle.

- Un facteur de transcription pionnier se fixe sur sa séquence cible de l'ADN, au sein de la région promotrice d'un gène.
- Ce facteur recrute un coactivateur à activité histone-acétyltransférase (HAT) qui modifie les nucléosomes environnants en acétylant les queues N-terminales d'histones sur des résidus spécifiques.
- Un complexe de remodelage de la famille SWI/SNF est recruté et assure la mobilisation ATP-dépendante des nucléosomes.
- Le relâchement local de la chromatine permet le démasquage de sites de liaison pour un nouveau facteur de transcription et le recrutement de co-activateurs transcriptionnels.
- Le complexe médiateur s'associe aux activateurs. Puis le complexe de pré-initiation s'assemble sur la boîte TATA, avant de mûrir vers l'état de complexe compétent pour l'élongation.

1.2.1.2. Mécanismes activateurs liés à l'acétylation locale des histones

Le promoteur d'un gène actif est caractérisé par un ensemble de marques épigénétiques associées à l'activation transcriptionnelle (acétylations de H3 et de H4, triméthylation H3K4 notamment) (**Figure 8A**). Comment se fait le lien entre modifications des histones et augmentation du taux de transcription d'une séquence génique ? Tout d'abord, la présence de groupements chimiques sur les histones peut influencer directement l'état chromatinien, en altérant l'accessibilité à l'ADN et les interactions inter- et intra-nucléosomales ([Strahl and Allis, 2000](#)). Ce modèle a été proposé pour l'acétylation des histones. Lorsque qu'un résidu lysine est acétylé, la charge positive portée par la lysine est neutralisée ([Kouzarides, 2007b](#)). Cette perte de charge diminuerait l'interaction avec l'ADN chargé négativement, donnant naissance à une chromatine plus flexible où l'ADN est plus accessible, et autorisant le recrutement de la machinerie transcriptionnelle ([Eberharder and Becker, 2002](#)). Un second mécanisme d'action, indirect, serait rendu possible par le recrutement sur les nucléosomes de protéines effectrices via des domaines protéiques reconnaissant spécifiquement les modifications des histones. Ces protéines effectrices peuvent ensuite altérer la structure de la chromatine ou favoriser le recrutement de l'ARN polymérase ([Taverna et al., 2007](#)). Le bromodomaine est un domaine protéique qui a été identifié comme interagissant spécifiquement avec les histones acétylées (**Figures 6C et 8A**). Le complexe de remodelage Swi/Snf est ainsi recruté sur les résidus d'histones acétylés via son bromodomaine ([Hassan et al., 2002](#)). Autre exemple de ce mécanisme, la protéine BRDT, spécifique des cellules germinales mâles, est capable de fixer H4 sous sa forme hyperacétylée via son premier bromodomaine ([Moriniere et al., 2009](#)). En association avec le complexe activateur pTEF β aux promoteurs hyperacétylés ([Gaucher et al., 2012](#)), elle permet la transcription de gènes méiotiques et post-méiotiques.

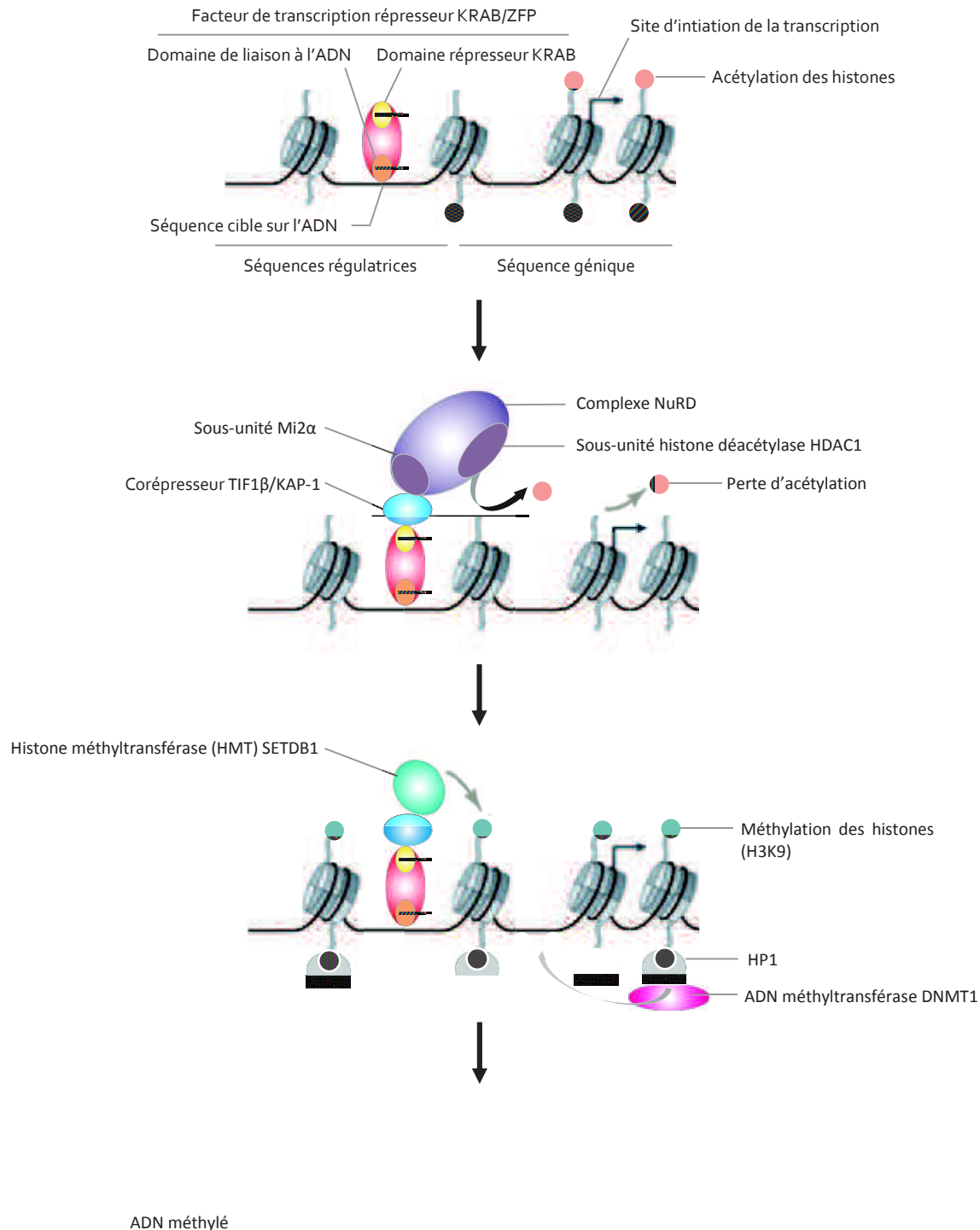
1.2.2. Mécanismes de répression transcriptionnelle

Dans la plupart des cas, l'inhibition de la transcription est en fait une désactivation. Malgré cela, des mécanismes de répression visent à bloquer la région promotrice et empêcher la fixation de l'ARN polymérase, soit en liant des protéines qui occupent physiquement l'espace dévolu au complexe de pré-initiation, soit en organisant la chromatine de manière ce que la région promotrice soit inaccessible (**Figure 10**).

1.2.2.1. Facteurs de transcription spécifiques à activité répressive

L'action des répresseurs transcriptionnels peut tout d'abord provoquer l'arrêt de toute nouvelle initiation de la transcription par déstabilisation de la machinerie transcriptionnelle. Ainsi, certains répresseurs occupent physiquement la région promotrice, empêchant toute fixation du CPI. D'autres inhibent l'assemblage de la machinerie transcriptionnelle. C'est le cas de la protéine Mot1/Btaf1 qui, une fois fixée en amont de la boîte TATA, catalyse la rupture ATP-dépendante du complexe TBP/ADN ([Darst et al., 2001](#) ; [Lee and Young, 1998](#)). Certains répresseurs entrent en interaction directe avec la machinerie transcriptionnelle et répriment son activité. Certains enfin agissent en masquant la surface d'activation d'un facteur activateur fixé à proximité, ou en entrant en compétition avec d'autres facteurs activateurs pour le site de liaison à l'ADN. Dans la famille Ets par exemple, le facteur de transcription Fev agit comme un répresseur transcriptionnel, en cumulant une activité de répression active (via son domaine carboxy-terminal) et un mode de répression passif par occupation des sites de liaison des facteurs Ets activateurs ([Maurer et al., 2003](#)).

Les répresseurs transcriptionnels peuvent également agir de manière indirecte en provoquant la condensation locale de la chromatine, par le recrutement d'HDAC, de HMT ou de complexes de remodelage. Snail, facteur de transcription se fixant sur les séquences « boîtes E » de l'ADN grâce à son domaine à doigts de zinc ([Verschuere et al., 1999](#)), agit par exemple en tant que répresseur transcriptionnel au cours de la transition épithélio-mésenchymateuse ([Nieto, 2002](#) ; [Peinado et al., 2007](#)). Il apparaît capable de recruter au promoteur d'E-cadhérine un complexe de répression composé de HDAC1, HDAC2 et Sin3A



ADN méthylé

Figure 10 : Mécanismes de répression transcriptionnelle

Il s'agit ici de la représentation des différentes étapes d'un modèle générique de répression transcriptionnelle, inspiré de la répression induite par les facteurs KRAB/ZFP. Inspiré de (Ryan et al., 1999 ; Schultz et al., 2001 ; Sripathy et al., 2006).

- Un facteur de transcription à rôle répresseur se fixe sur sa séquence cible de l'ADN, au sein de la région promotrice d'un gène, grâce à son domaine en doigts de zinc.
- Via le coactivateur TIFβ (ou KAP-1) qui interagit avec Mi2α, le complexe répresseur NuRD est recruté. Son activité HDAC entraîne la perte des modifications post-traductionnelles activatrices des queues d'histones, acétylations principalement.
- L'histone méthyltransférase qui s'associe au corépresseur permet d'apposer de nouvelles marques épigénétiques répressives sur les nucléosomes (di- et tri-méthylations d'H3K9).
- Le chromodomaine d'HP1 lui permet de se fixer sur les H3K9 méthylées. HP1 agit alors comme une plate-forme, permettant notamment le recrutement d'une ADN méthyltransférase. On aboutit à une répression transcriptionnelle forte par méthylation de l'ADN et compaction locale de type "hétérochromatine".

(Peinado et al., 2004). Les récepteurs nucléaires, en l'absence de leurs ligands, s'associent à des corépresseurs tels que SMRT NcoR qui font le lien avec des HDAC (Chen and Evans, 1995 ; Horlein et al., 1995). Les facteurs de transcription à doigts de zinc de la famille KRAB-ZFP (Krüppel Associated Box – Zinc Finger Protein) (**Figure 10**) interagissent avec le corépresseur KAP-1 (KRAB-ZFP Associated Protein 1). Ce corépresseur se lie à la sous-unité Mizα du complexe répresseur NuRD/HDAC1 (Schultz et al., 2001) et à l'HMT SETDB1, conduisant au recrutement de la protéine HP1 (Heterochromatin Protein 1) au promoteur des gènes cibles (Nielsen et al., 1999 ; Ryan et al., 1999) et à une répression transcriptionnelle forte et stable (Ayyanathan et al., 2003 ; Sripathy et al., 2006). A terme, l'expression d'un gène peut être verrouillée par hétérochromatinisation de la région (Groner et al., 2010).

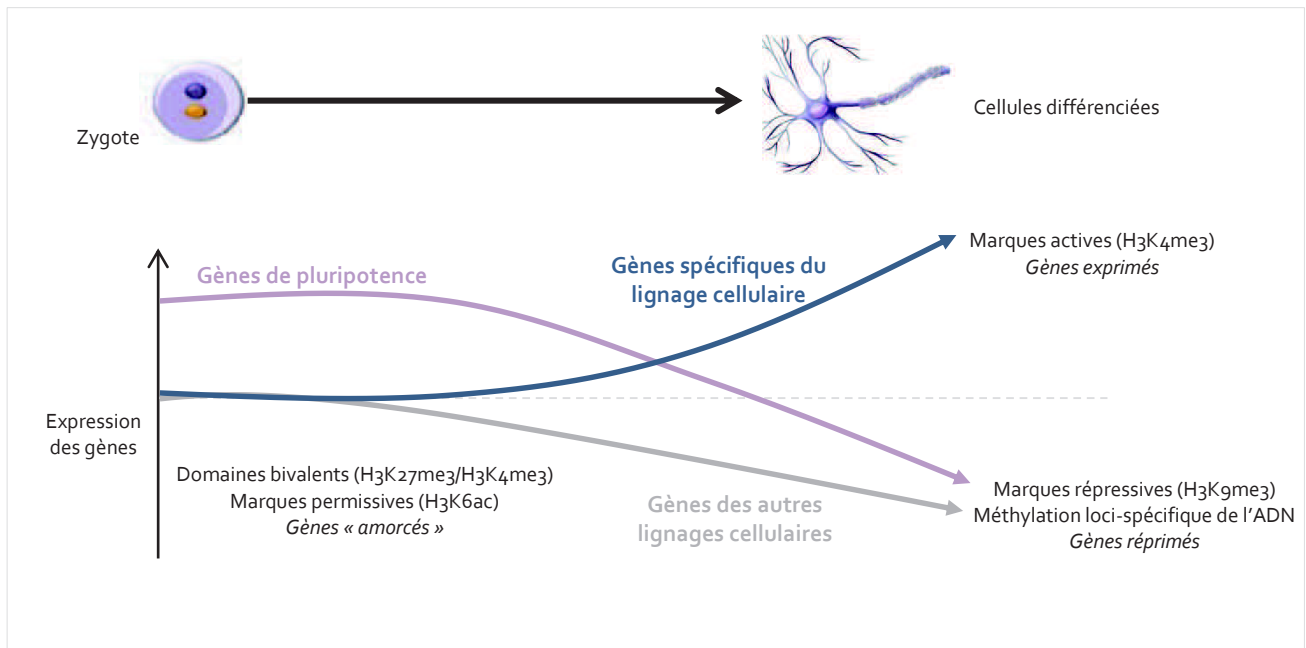
1.2.2.2. Mécanismes répresseurs liés à la méthylation des histones

Les modifications post-traductionnelles des histones peuvent être reconnues par des répresseurs, recrutés sur les nucléosomes et agissant directement sur les séquences promotrices des gènes proches. Ainsi, la méthylation des lysines est reconnue par les domaines chromodomaines, Tudor, MBT et PHD présents chez de nombreux répresseurs transcriptionnels (**Figure 6C**) (James and Frye, 2013). L'H3K27 méthylée est par exemple reconnue par la protéine à chromodomaine Polycomb qui intervient dans la répression de gènes durant le développement (Cao et al., 2002). L'H3-K9 di- ou tri-méthylée est quant à elle reconnue par les protéines de la famille HP1 (heterochromatin protein 1) (Maison and Almouzni, 2004). HP1 est surtout connue pour être présente au niveau de l'hétérochromatine, mais est également retrouvée au niveau de l'euchromatine où sa présence est souvent corrélée à une répression de gènes (Fanti et al., 2003 ; Grewal and Moazed, 2003 ; Hwang et al., 2001). Les protéines HP1 associées aux H3K9 méthylées pourraient jouer le rôle de plateforme (Grewal and Jia, 2007) pour le recrutement de protéines telles que les DNMT et l'HMT Suv39h1 (Bachman et al., 2001 ; Fuks, 2003 ; Lehnertz et al., 2003), l'ensemble permettant d'établir une structure chromatinienne répressive de type hétérochromatine (**Figure 10**).

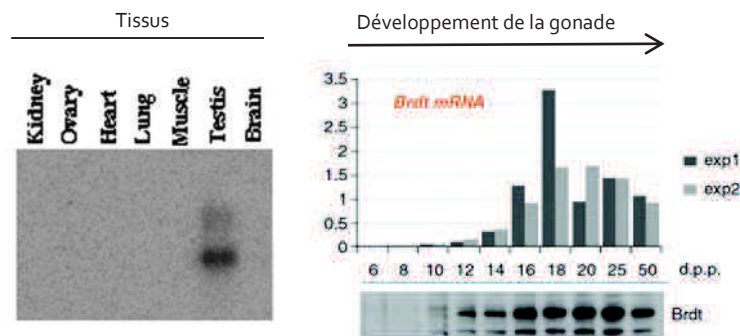
1.2.2.3. Mécanismes répresseurs liés à la méthylation locale de l'ADN

La méthylation de l'ADN au niveau d'une séquence régulatrice peut réprimer la transcription en empêchant la fixation de facteurs de transcription sur des éléments de réponse possédant des sites CpG (Costello et al., 2000 ; Watt and Molloy, 1988). Ainsi les activateurs transcriptionnels Sp1 et CREB ne se lient plus à leurs cibles si elles sont méthylées. De nombreux sites de méthylation sont retrouvés à l'intérieur même des unités de transcription (Klose and Bird, 2006) et perturberaient l'étape d'allongement de l'ARN par l'ARN polymérase II (Lorincz et al., 2004). Mais les conséquences de la méthylation peuvent aussi être indirectes et nécessiter l'intervention d'interprètes (**Figure 5C**). Les protéines de la famille MBP (Methyl-CpG Binding Protein) reconnaissent spécifiquement les îlots CpG méthylés par un motif conservé de 70 acides aminés (Methyl CpG Binding Domain, MBD) (Defossez and Stancheva, 2011). MeCP2, répresseur fort de l'expression de certains gènes neuronaux, assure sa fonction en interférant avec la machinerie transcriptionnelle par séquestration de TFIIB et en s'associant au complexe répresseur Sin3A via son domaine TRD (Transcriptional Repression Domain) (Nuber et al., 2005). MBD2 recrute quant à lui le complexe répresseur NuRD sur les promoteurs méthylés (Hutchins et al., 2002). Les protéines MBP conduisent également à l'apposition de nouvelles marques épigénétiques. Plusieurs MBP s'associent à des complexes histones déacétylases et MBD1 recrute l'H3K9-méthyltransférase SETDB1 (Sarraf and Stancheva, 2004). Globalement ces protéines « lecteurs » de la méthylation de l'ADN vont agir et/ou recruter des facteurs afin d'apporter des marques épigénétiques et/ou une activité de remodelage rendant la région génomique ciblée inapte à la transcription. La méthylation de l'ADN fait donc partie intégrante des mécanismes de répression transcriptionnelle, et on a montré que des gènes inactifs pouvaient produire des transcrits suite à la déméthylation des CpG induite par la 5-azacytidine (Chiurazzi et al., 1998).

A.



B.



C.

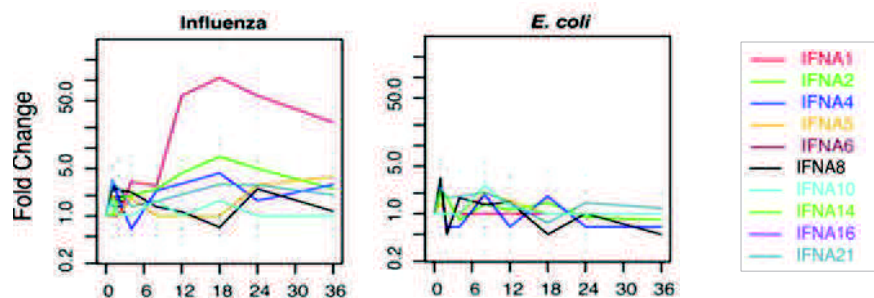


Figure 11 : Profils d'expression des gènes, stabilité de l'identité cellulaire et adaptabilité

A. Le paysage épigénétique conditionne les expressions géniques au cours du développement. Au cours de la transition entre le zygote et les cellules différenciées, la méthylation de l'ADN et les modifications d'histone régulent l'expression des gènes de pluripotence et des gènes spécifiques des lignées cellulaires. Chez le zygote, des gènes développementaux clés sont amorcés via des marques bivalentes. Au cours de la différenciation, les gènes spécifiques de la lignée sont activés par les marques épigénétiques, établissant l'identité cellulaire, alors que les gènes spécifiques de lignées alternatives et les gènes de pluripotence sont réprimés définitivement. *Inspiré de (Berdasco and Esteller, 2010).*

B. Le gène *BRDT* n'est exprimé que dans le testicule. Son expression est activée au cours de la différenciation des cellules germinales. *Tiré de (Pivot-Pajot et al., 2003) et (Gaucher et al., 2012).*

C. Les expressions géniques sont variables et assurent l'adaptabilité d'une cellule aux stimuli. Patron d'expression des gènes *IFN-α* dans une cellule dendritique au cours du temps après infection par un pathogène obtenu par analyse sur puce. *Tiré de (Zilliox et al., 2006).*

Cette partie a tenté d'illustrer quelques mécanismes généraux de régulation transcriptionnelle. On voit que les protéines régulatrices agissent rarement seules : la régulation est souvent assurée par de larges complexes multiprotéiques associant facteurs de transcription, coactivateurs et/ou corépresseurs, établis grâce aux interactions ADN/protéines et protéines/protéines. Il faut signaler que certaines protéines régulatrices ne sont pas intrinsèquement activateur ou répresseur, mais peuvent être recrutées dans plusieurs complexes dont l'activité finale (activation ou répression) dépendra de l'assemblage différentiel et combinatoire des différents composants (Goodman and Smolik, 2000). Différents complexes peuvent agir en concomitance ou de façon séquentielle lors de l'activation ou de la répression transcriptionnelle : ils peuvent être recrutés suite à la modification apportée à la chromatine par le complexe les précédant (Figures 9 et 10).

1.3. Répertoire restreint d'expressions géniques et identité cellulaire

Au sein du corps humain, les cellules différenciées possèdent des caractéristiques structurales et fonctionnelles souvent uniques. Or les expériences de clonage par transfert de noyau menées dès la fin des années 1950 (Gurdon et al., 1958) ont démontré que toutes les instructions nécessaires à la constitution de l'organisme entier sont présentes dans chaque cellule somatique. Ainsi, une information génétique identique conduit à l'élaboration de phénotypes cellulaires différents. Ceci peut être expliqué par l'activation sélective par chaque cellule d'un répertoire restreint de gènes. Alors que le génome humain contient plus de 20 000 gènes codant des protéines, chaque type cellulaire n'en exprime qu'un sous-ensemble. Ainsi, une cellule à mucus de la muqueuse bronchique n'exprime pas le même sous-ensemble de gènes qu'un neurone moteur, et réprime à l'inverse l'expression de gènes qui seront activés dans les cellules musculaires lisses. L'expression d'un sous-groupe de gènes est permise par les phénomènes épigénétiques d'une part et par la combinaison de facteurs de transcription présents et activés dans la cellule à un moment donné d'autre part (Figure 12).

Les phénomènes épigénétiques interviennent dans le contrôle des expressions géniques dès les premières étapes du développement embryonnaire. Ils sont par exemple impliqués dans l'inactivation d'un chromosome X chez la femelle, lorsque l'hétérochromatinisation de l'un des deux gonosomes réprime l'expression des gènes qu'il porte, évitant le surdosage dû à la présence de deux allèles chez la femelle XX (Ohno et al., 1959). La méthylation de l'ADN joue un rôle fondamental au cours des étapes précoces du développement dans la mise en place de l'empreinte parentale, c'est-à-dire lorsqu'un allèle doit être exprimé différemment selon l'origine parentale (Delaval and Feil, 2004 ; Morison et al., 2005 ; Wolffe and Matzke, 1999). Les gènes soumis à empreinte maternelle possèdent ainsi une région ICR (Imprinting Control Region) ou DMR (Differentially Methylated Region) qui est fortement méthylée sur l'allèle paternel, réprimant son expression (Lewis et al., 2004). Les mécanismes épigénétiques établissent également une base fondamentale pour le maintien de la totipotence des cellules souches embryonnaires (ES) (Berdasco and Esteller, 2010) (Figure 11A). Celles-ci possèdent des domaines génomiques bivalents qui portent des modifications d'histones de fonctions opposées. Les formes méthylées de H3K27 (répressive) et H3K4 (activatrice) notamment coexistent sur les promoteurs de facteurs de transcription. Lorsque ces cellules ES se différencient, et en fonction des signaux reçus, ces domaines bivalents ne conservent qu'une seule de ces modifications, conduisant à l'expression ou à la répression du facteur de transcription – et donc à la mise en place ou non d'un programme génétique propre au lignage cellulaire emprunté (Dillon, 2012 ; Kouzarides, 2007b). Lors de la différenciation cellulaire, la méthylation des promoteurs joue également un rôle central dans la répression des gènes spécifiques d'autres lignages cellulaires qui doivent être réduits au silence transcriptionnel, ce qui participe à l'élaboration d'un programme génétique propre à chaque cellule. Chez l'embryon humain, dans les cellules de la lignée érythrocytaire, les promoteurs des gènes codant les ϵ -globines formant l'hémoglobine embryonnaire sont déméthylés (Groudine and Weintraub, 1981 ; van der Ploeg and Flavell, 1980). Quand l'hémoglobine embryonnaire doit laisser place à la forme adulte, on observe une méthylation des promoteurs des ϵ -globines associée à un silence transcriptionnel, ainsi qu'une déméthylation des promoteurs des gènes codant les γ -globines qui sont alors exprimés (Mavilio et al., 1983).

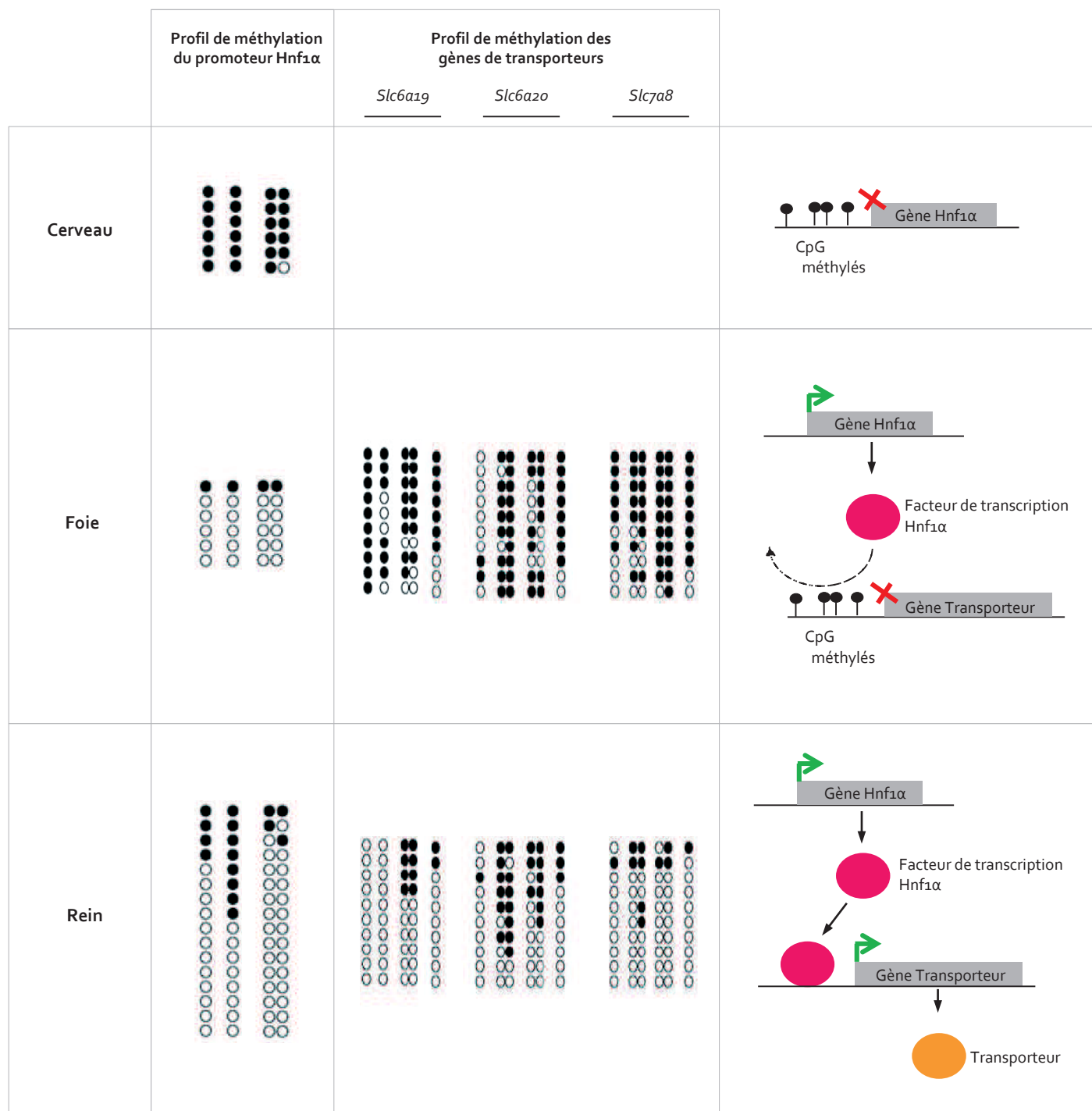


Figure 12 : Facteurs de transcription et marques épigénétiques se combinent pour établir les patrons d'expression des gènes

La méthylation de l'ADN et le facteur de transcription *Hnf1α* (Hepatocyte nuclear factor 1) se combinent pour contrôler l'expression des gènes codant des transporteurs d'acides aminés dans les cellules du rein. Le gène *Hnf1α* et les gènes codant les transporteurs possèdent des régions différemment méthylées selon les tissus (T-DMR) à proximité de leur signal de début de transcription (TSS). Dans le cerveau, l'hyperméthylation de *Hnf1α* empêche la production du facteur de transcription. Dans le foie et le rein, le promoteur de *Hnf1α* est hypométhylé et le facteur de transcription est produit. Dans le foie, l'hyperméthylation des promoteurs de transporteurs empêche la fixation de la protéine *Hnf1α* sur l'ADN. Dans les cellules rénales, la fixation du facteur de transcription sur le promoteur déméthylé autorise l'expression des transporteurs d'acides aminés.

Adapté de (Kikuchi et al., 2010).

Le contrôle des expressions géniques par les facteurs de transcription est également un des fondements de la différenciation et de la morphogenèse. En se liant à de nombreux promoteurs présentant sa séquence de liaison spécifique, un unique facteur de transcription peut déclencher l'expression d'un répertoire d'expressions géniques défini. La présence ou l'absence des facteurs de transcription, leurs activités et leur fonctionnement combinatoire, vont ainsi diriger des programmes d'expression particuliers conditionnant l'identité cellulaire. Les facteurs de transcription sont également des facteurs clés dans l'adaptation des cellules aux variations de l'environnement et la régulation temporelle des expressions. Vers les facteurs de transcription convergent en effet de nombreuses voies de signalisation. Suite à un stimulus donné, l'activité des facteurs de transcription peut être modulée, et conduire à la régulation des réseaux d'expressions géniques, impactant directement le comportement cellulaire et la réponse de la cellule aux variations de son milieu (Tootle and Rebay, 2005).

Les différentes cellules du corps humain possèdent donc des répertoires d'expression distincts, participant à l'établissement d'une identité stable, mais adaptable face aux variations de l'environnement. Le corollaire est qu'un même gène peut ne pas être exprimé dans toutes les cellules de l'organisme. Dans certains cas, son expression peut même être restreinte à un seul organe ou un seul type cellulaire, ou être spécifique d'une condition donnée (**Figure 12**). Pour chaque gène, on peut alors définir un patron spatio-temporel d'expression, ce patron étant principalement dirigé par les séquences régulatrices de ce gène. Ainsi, on distingue les promoteurs constitutifs, actifs continuellement, qui contrôleront l'expression de gènes exprimés dans l'ensemble des cellules de l'organisme, et les promoteurs régulés par des facteurs de transcription spécifiques permettant la transcription des gènes qu'ils contrôlent uniquement dans certains environnements cellulaires particuliers ou en réponse à certains stimuli. Les gènes peuvent ainsi être régulés de manière spatiale et donner naissance à des profils d'expression tissu-spécifiques ou type cellulaire-spécifiques. Parmi les gènes dont l'expression est restreinte à un tissu donné, on retrouve en particulier de nombreux gènes spécifiques du testicule et de la lignée germinale mâle (Rousseaux, 2013) (voir paragraphe 3.3), comme par exemple le gène *BRDT* qui ne s'exprime que dans les cellules germinales post-méiotiques (**Figure 11B**) (Gaucher et al., 2012 ; Pivot-Pajot et al., 2003). Les gènes peuvent également être induits de manière temporelle, comme les gènes des cyclines dont l'expression contrôle la progression dans le cycle cellulaire, les gènes codant les protéines de choc thermique (Heat Shock Proteins - HSP) exprimées consécutivement à un stress thermique, ou encore les gènes interférons *IFN* activés suite à une infection par un agent pathogène (**Figure 11C**) (Zilliox et al., 2006).

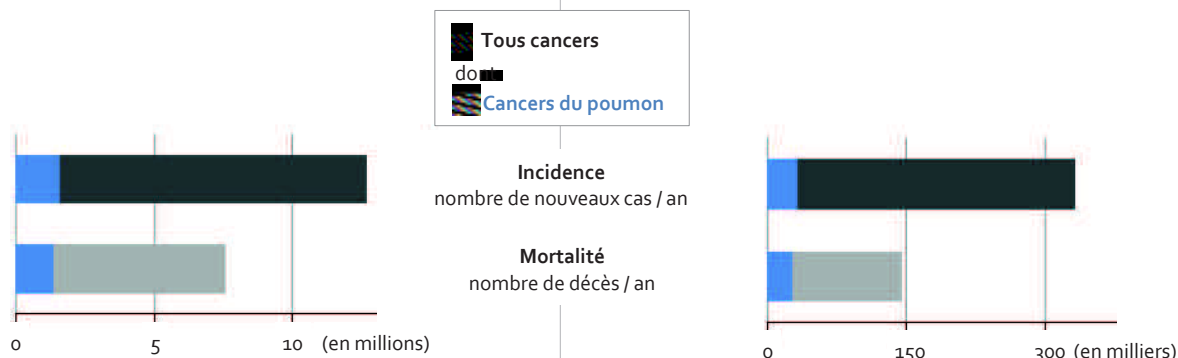
Le contrôle des expressions géniques est un processus fondamental de la vie, base de la différenciation et de la morphogenèse, mais également de la versatilité et de l'adaptabilité d'un organisme. Depuis les premières étapes du développement, les identités cellulaires sont établies grâce à la mise en place d'un répertoire d'expressions géniques particulier qui va conditionner le devenir de chaque cellule et la stabilité de l'identité cellulaire. En conséquence, tout dérèglement des mécanismes de régulation de l'expression des gènes peut avoir des conséquences drastiques, non seulement pour la cellule mais également pour l'organisme entier. Etant donné l'importance du maintien de l'identité cellulaire, beaucoup d'espairs reposent sur la contribution que pourrait apporter la compréhension de l'expression différentielle des gènes à la lutte contre certaines maladies, depuis le vieillissement jusqu'au cancer (Bentley, 1999). Les cellules cancéreuses perdent en effet la capacité de maintenir une identité correcte et exhibent un patron aberrant de gènes exprimés. C'est la notion qui sera développée dans la troisième partie de cette introduction, après une présentation de notre modèle d'étude, le cancer du poumon.

Monde

France



A.

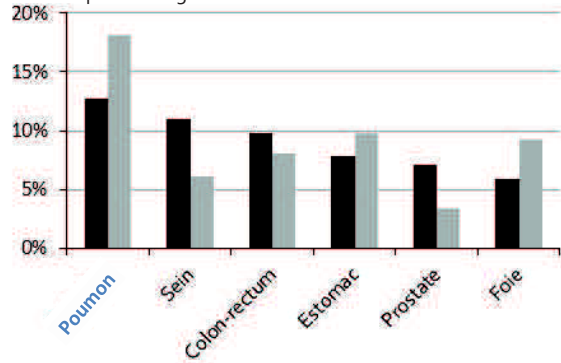


B.

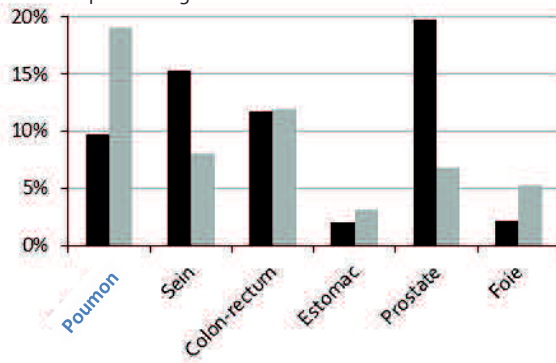
Contribution des différents types de cancers dans

incidence
mortalité

en pourcentage de l'ensemble des cancers



en pourcentage de l'ensemble des cancers



C.

Analyse selon le sexe

incidence
mortalité

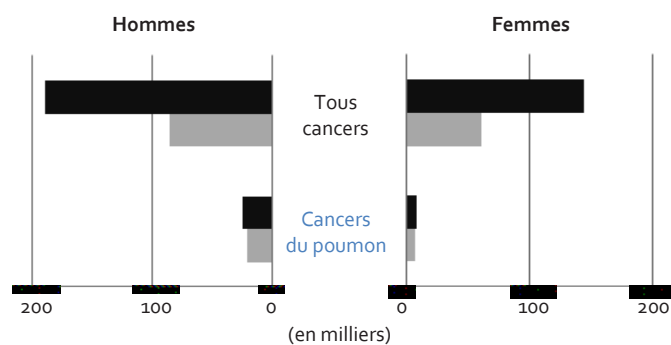
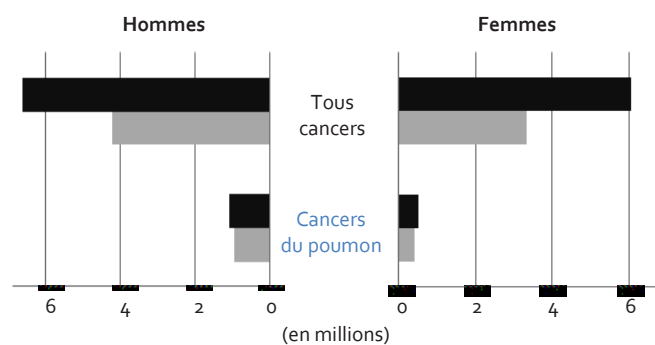


Figure 13 : Données chiffrées sur l'incidence et la mortalité des cancers du poumon, dans le monde (colonne de gauche) et en France (colonne de droite)

Graphiques construits à partir des données chiffrées disponibles dans la base de données GLOBOCAN 2008.

2. LE CANCER DU POUMON

Les cancers constituent une cause majeure de décès dans le monde. En 2008, 12,7 millions de nouveaux cas ont été recensés ; plus de 7,5 millions de décès sont survenus, ce qui représente environ 13% de la mortalité mondiale (Ferlay et al., 2010) (Figure 13A). La France n'est pas épargnée : plus de 357 000 nouveaux cas de cancer ont été diagnostiqués en 2010. Avec 146 000 décès estimés en 2010, les pathologies tumorales constituent la première cause de mortalité en France (INCa, 2010) (Figure 13A). A ce bilan humain s'ajoutent les coûts socio-économiques du cancer. Les chirurgies lourdes, les chimio- et radio-thérapies en font l'une des maladies les plus coûteuses, si bien que les dépenses de soins et de prévention sont estimées à plus de 12 milliards d'euros par an en France (Almaric, 2007). La maladie a également des effets dévastateurs sur le plan sociétal : choc de l'annonce sur le malade et sa famille, dégradation de la qualité de vie, perte d'utilité sociale, deuil. En conséquence, la lutte contre les cancers est devenue un enjeu majeur de santé publique, donnant lieu à d'intenses efforts de recherche pour tenter de mieux comprendre, de mieux détecter et de mieux soigner cette maladie.

2.1. Epidémiologie et étiologie du cancer du poumon

2.1.1. Un cancer fréquent et de pronostic sombre

2.1.1.1. Incidence : Nombre de nouveaux cas détectés chaque année

Représentant 12% de l'ensemble des cancers dans le monde, le cancer du poumon est le cancer le plus fréquent sur la planète (Figure 13B). En France, le cancer du poumon représente le 4^{ème} type de cancer le plus diagnostiqué ; avec 37 000 nouveaux cas estimés en 2010, il constitue presque 10% du total des cancers recensés (INCa, 2010 ; INCa, 2012) (Figures 13B et 14A). Les cancers dans leur ensemble touchent davantage les individus de sexe masculin, mais les cancers du poumon présentent un sex-ratio particulièrement déséquilibré avec 84 % des cas diagnostiqués chez des hommes sur la période 1889-2007 (Grosclaude et al., 2013) (Figure 13C). Par ailleurs, l'âge moyen du diagnostic du cancer du poumon était estimé à 65 ans en 2005 (Belot et al., 2008) et il est rarement diagnostiqué avant 50 ans (moins de 5% des cas).

Le nombre annuel de cas de cancers du poumon recensés a connu une hausse importante jusqu'à la fin des années 1970 et continue d'augmenter aujourd'hui, bien que de manière moins marquée (Figure 14B). Cette augmentation continue de l'incidence est en premier lieu le reflet de la consommation de tabac, dont le pic a été atteint en France en 1983 (Guerin and Hill, 2010) mais il faut également prendre en compte l'amélioration des techniques de diagnostic dans l'augmentation du nombre de nouveaux cas recensés (Welsh et al., 2004). Par ailleurs, il est intéressant de noter les évolutions différentes de l'incidence entre les sexes : elle tend à se stabiliser chez les hommes depuis 2000, alors que le nombre de cancers du poumon diagnostiqués chaque année chez les femmes ne cesse d'augmenter. Celui-ci a ainsi triplé au cours de ces 20 dernières années (INCa, 2010 ; INCa, 2012), reflétant principalement l'augmentation du tabagisme féminin.

2.1.1.2. Pronostic : Données de survie à 5 ans

Le pronostic des différents cancers, qui varie considérablement selon la localisation tumorale, peut être approché par l'évaluation du taux de survie à 5 ans après le diagnostic (Figure 14A). Celui-ci dépasse 90 % pour les cancers du testicule ou de la thyroïde par exemple, mais est estimé à 14,5% pour les cancers du poumon (Berrino et al., 2007). Ces données font du cancer du poumon l'un des cancers au pronostic le plus sombre avec le mésothéliome pleural (7%) ou les tumeurs du pancréas (5%). De plus, aucune amélioration globale de la survie n'a été observée en France entre 1989-1991 et 2001-2004 (Grosclaude et al., 2013), malgré

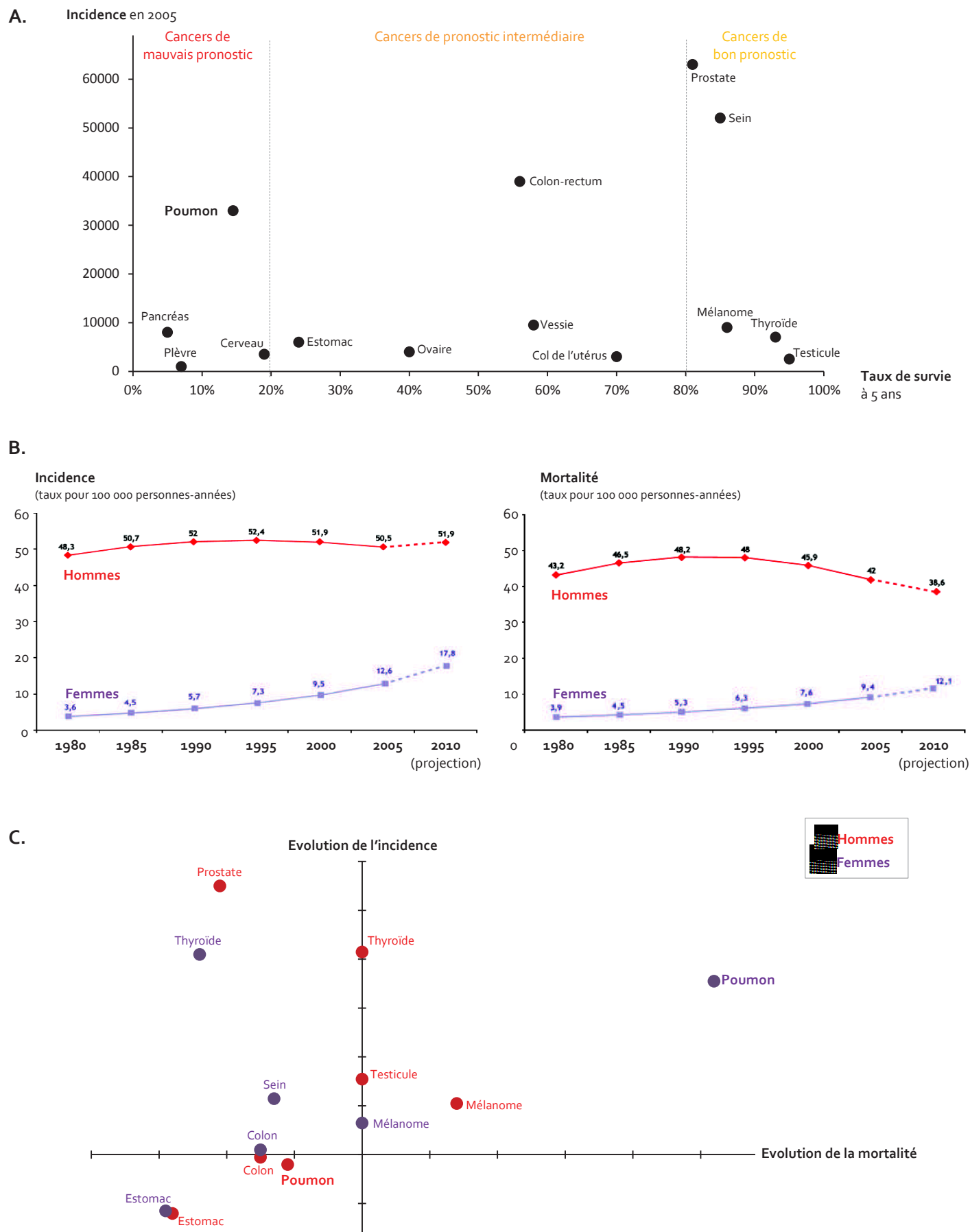


Figure 14 : Données chiffrées sur la survie, l'incidence et la mortalité des cancer du poumon

A. Survie des patients français en fonction de la localisation tumorale.

B. Evolutions de l'incidence et de la mortalité du cancer du poumon en France durant les dernières années.

C. Evolutions conjointes de l'incidence et de la mortalité des différents cancers en France au cours des dix dernières années
Adapté de (INCa, 2010 ; INCa, 2012).

le développement de techniques d'imagerie et de traitements innovants. La faible survie des patients atteints de cancers du poumon pourrait notamment s'expliquer par le diagnostic tardif des tumeurs pulmonaires, dont les symptômes n'apparaissent qu'à un stade avancé de la maladie (**voir paragraphe 2.4.1**). On peut noter que la survie à 5 ans est légèrement supérieure chez la femme (18 % contre 13% chez l'homme) ([Grosclaude et al., 2013](#)) : ce meilleur pronostic chez les femmes, déjà observé ([Patel, 2005](#)), pourrait s'expliquer par une plus grande attention portée par les femmes sur leurs éventuels symptômes (**voir paragraphe 2.4.1**), conduisant à des diagnostics plus précoces.

2.1.1.3. Mortalité : Nombre de décès par an

Conséquence de l'incidence élevée et du pronostic défavorable du cancer du poumon, celui-ci constitue la première cause de mortalité par cancer dans le monde, responsable de plus d'un million de morts autour du globe en l'an 2000 ([Ferlay et al., 2010](#)) (**Figure 13B**), devenant ainsi le cancer « le plus préoccupant en termes de fréquence et de pronostic » selon l'Organisation Mondiale de la Santé (OMS). C'est également la première cause de mortalité par cancer en France (près de 20% de l'ensemble des décès par cancers), avec 28 700 décès estimés en 2010 ([Grosclaude et al., 2013](#)) (**Figure 13B**), dont 73% chez l'homme et plus de 40% de façon prématurée (avant 65 ans).

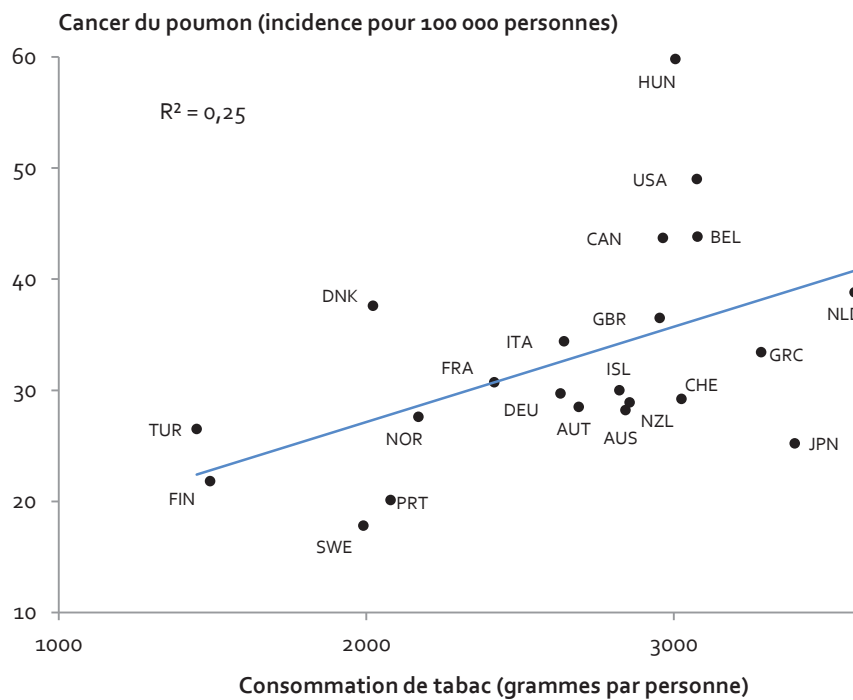
Comme pour l'incidence, la mortalité par cancer du poumon en France est en recul chez l'homme depuis les années 1990, alors qu'elle continue à augmenter drastiquement chez la femme : le taux de mortalité chez la femme par cancer du poumon a plus que doublé au cours des 20 dernières années ([Grosclaude et al., 2013](#)) (**Figures 14B et C**). Cette évolution contrastée entre les deux sexes est là encore le reflet de la diminution importante de la consommation de cigarettes chez les hommes et de l'entrée plus récente des femmes dans le tabagisme (à la fin des années 1960). L'augmentation du nombre de fumeuses devrait avoir pour conséquence de tripler le nombre de décès par cancer chez les femmes d'ici 2020.

2.1.2. Un cancer imputable en majorité à la consommation de tabac

2.1.2.1. Le tabac et les carcinogènes de la fumée de cigarette

Les premières observations épidémiologiques sur les cancers du poumon datent du 19^{ème} siècle, époque à laquelle on note une fréquence élevée de cancers du poumon chez les mineurs exposés au radon ([Pirozynski, 2006](#)). Malgré cela, le cancer du poumon était une maladie excessivement rare : lorsque Adler publie en 1912 un livre recensant les cas de cancers du poumon relatés dans la littérature médicale mondiale, il n'en comptabilise que 374 ([Adler, 1912 ; Spiro and Silvestri, 2005](#)). A partir de la fin des années 1920, on observe une explosion du nombre de cas, rapidement imputée à un nouvel acteur : le tabac, dont la consommation s'est intensifiée à partir du début du 20^{ème} siècle (**Figure 15B**) ([Spiro and Silvestri, 2005](#)). Dès 1929, le médecin allemand Fritz Lickint reconnaît le lien entre tabac et cancer du poumon ([Witschi, 2001](#)) et la première preuve épidémiologique solide de l'implication du tabagisme est fournie dans les années 1950 ([Doll and Hill, 1956](#)). Les données récentes ne cessent de démontrer la corrélation entre consommation de tabac et cancer du poumon, aussi bien en terme d'incidence (**Figure 15A**) qu'en terme de mortalité (**Figure 15B**). De manière intéressante, on observe toujours un décalage de 10 à 30 ans entre consommation de tabac et cancer du poumon ([Spiro and Silvestri, 2005](#)). Il est ainsi clairement établi aujourd'hui que la cause la plus fréquente de cancer du poumon est l'exposition chronique à la fumée de cigarette : 85% des tumeurs pulmonaires sont diagnostiquées chez les fumeurs et le tabac entraînerait directement 71% des décès par cancer du poumon dans le monde ([Hecht, 2003](#)). Le risque lié au tabac augmente avec le volume quotidien de la consommation, sa durée, l'intensité de l'inhalation de la fumée et varie selon le mode de consommation (présence ou non d'un filtre sur les cigarettes, type de tabac, teneur en goudrons). Le tabagisme passif a également été reconnu comme facteur de risque pour le cancer du poumon ([WHO, 2002](#)) : il accroîtrait le risque de survenue d'un cancer du poumon de 25 % ([Hackshaw et al., 1997](#)). En conséquence, le tabagisme a constitué la cible majeure

A.



B.

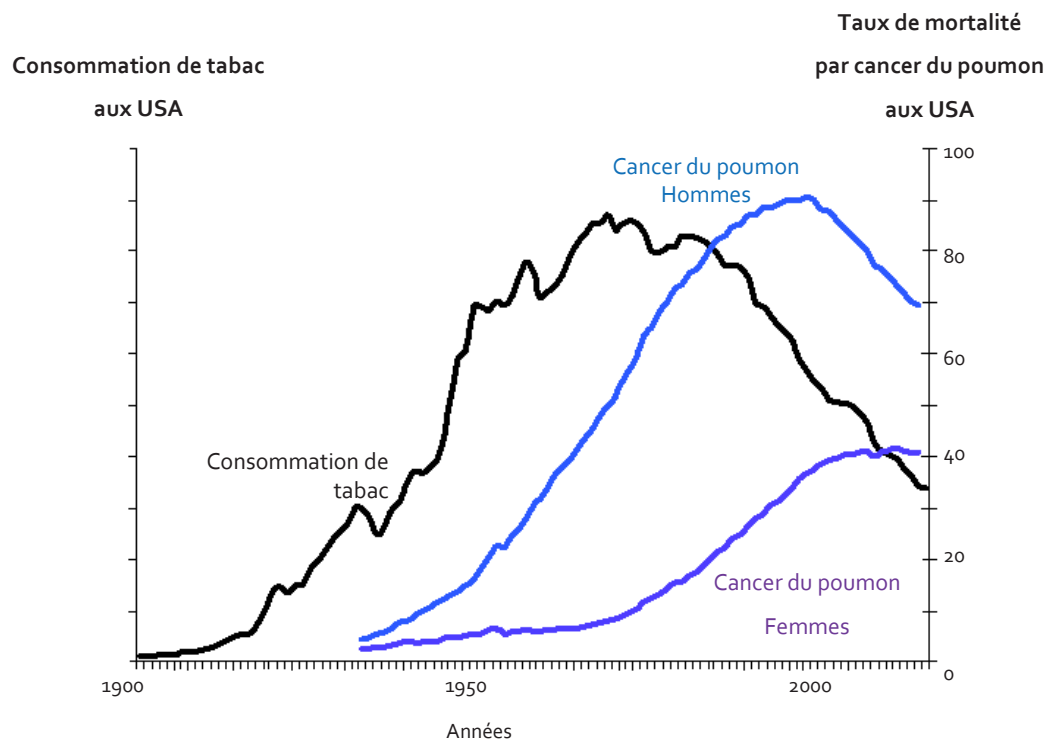


Figure 15 : Implication de la consommation de tabac dans les cancers pulmonaires

A. Corrélation entre consommation de tabac et incidence du cancer du poumon dans les pays développés.

Adapté de « Panorama de la santé 2009: Les indicateurs de l'OCDE ».

B. Evolutions temporelles de la consommation de tabac et du taux de mortalité par cancer du poumon aux USA.

Traduit de « Cancer Statistics 2009 », une présentation de the American Cancer Society.

des campagnes de prévention mises en place depuis les années 1980. Ainsi, la réduction de la consommation de tabac serait à l'origine de la baisse d'incidence et de mortalité amorcée chez les hommes vers 1990-2000 (Figures 14B).

Le tabagisme provoque l'exposition chronique à de nombreuses substances carcinogènes (Figure 16A) (Hecht, 2003), telles que les hydrocarbures aromatiques polycycliques (HAP) ou les nitrosamines présents dans la fumée de cigarette. Ces carcinogènes sont à l'origine d'altérations diverses dans les cellules pulmonaires (Figure 16C), souvent après activation métabolique (par le cytochrome p450 par exemple) (Luch and Baird, 2005). Leur mode d'action principal implique la formation d'adduits à l'acide désoxyribonucléique (ADN), des groupements chimiques réactifs se liant de manière covalente aux bases azotées (Figure 16B). Non réparés, ces adduits aboutissent à des mésappariements lors de la réplication de l'ADN puis à des mutations permanentes pouvant affecter des gènes clés (Figure 16C) (voir paragraphe 2.3.1) (Sturla, 2007). La nicotine et certains carcinogènes peuvent également se lier à des récepteurs cellulaires et affecter des voies de signalisation clés (comme la voie PI3K-AKT-mTOR (Figure 23)) pour la régulation de l'apoptose, la prolifération ou la survie des cellules pulmonaires (voir paragraphe 2.3.2) (Hecht, 2003). Enfin, on considère que le tabagisme pourrait favoriser la carcinogenèse bronchique en maintenant un état chronique d'inflammation au sein du poumon, ce microenvironnement inflammatoire jouant un rôle important de promotion de la croissance tumorale (Hecht, 2012) (voir paragraphe 2.2.1.3).

2.1.2.2. Autres facteurs environnementaux carcinogènes

Le cancer du poumon n'est pas exclusivement lié au tabac. D'autres facteurs exogènes comme les radiations, l'amiante ou la pollution atmosphérique expliquent le développement de cancers du poumon chez les non-fumeurs, qui représentent environ 15 % des patients atteints de tumeurs pulmonaires (Thun et al., 2008). Par ailleurs, il est intéressant de préciser que ces différents facteurs ont des effets sur la carcinogenèse pulmonaire distincts de ceux mis en cause dans le cas du tabac (Hubaux et al., 2012).

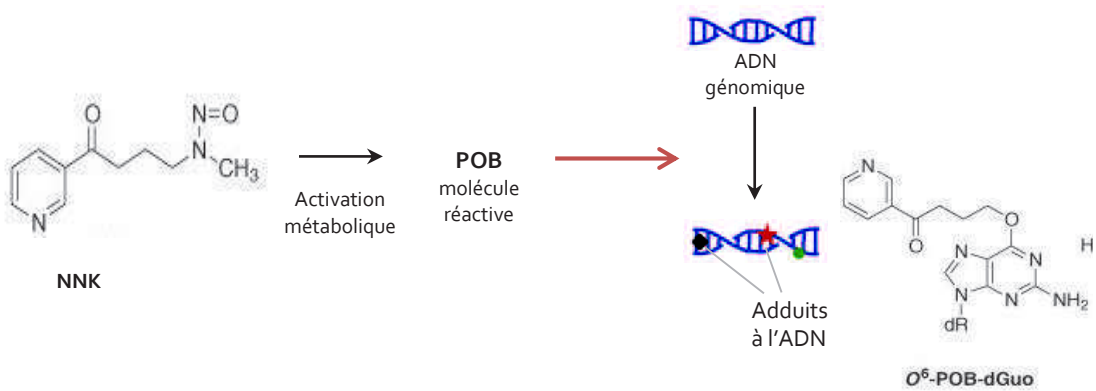
Le lien avec le radon a été établi pour la première fois dans les mines d'argent de l'Erzgebirge ; les mineurs étaient atteints d'une quantité exceptionnelle de maladies respiratoires, finalement reconnues comme cancers du poumon dans les années 1870 (Pirozynski, 2006). Le radon, gaz radioactif, est issu de la désintégration de l'uranium présent dans les roches. Une fois dans les tissus pulmonaires, les produits de désintégration du radon génèrent des particules alpha qui endommagent l'ADN des cellules de l'épithélium bronchique (Hubaux et al., 2012). Sur la double base des études épidémiologiques et expérimentales, le radon a été reconnu en 1987 cancérigène pulmonaire pour l'homme par le CIRC, même à des taux couramment rencontrés dans l'air intérieur des maisons. L'exposition au radon serait responsable en France de 13 % des décès par cancer du poumon, soit environ 3 350 morts chaque année (Catelinois et al., 2006).

Autre facteur critique dans l'apparition de cancers pulmonaires, l'amiante est une matière minérale fibreuse qui a été massivement utilisée en France jusqu'à son interdiction en 1997 ((96-1133, 1996), notamment pour l'isolation thermique des bâtiments. Les fibres d'amiante exercent leur effet cancérigène par l'intermédiaire d'interactions physiques directes avec les cellules pulmonaires et d'effets indirects sur les cellules immunitaires. Les fibres d'amiante inhalées ne peuvent pas être éliminées par phagocytose lorsque leur taille excède 5 µm ; elles entraînent alors la production d'espèces réactives de l'oxygène causant des dommages à l'ADN. De plus, les fibres d'amiante adhèrent à la chromatine, d'où d'importants défauts mitotiques tels que des ségrégations chromosomiques aberrantes (Heintz et al., 2010). Cette cascade d'évènements moléculaires initiée par l'amiante aboutit à l'inflammation, à la fibrose pulmonaire voire à la carcinogenèse. Principalement mise en cause dans les tumeurs de la plèvre, l'amiante est en réalité responsable de deux fois plus de morts par cancers du poumon (2000 à 3000 en France) que par mésothéliome pleural (Imbernon, 2003).

A.

Carcinogènes impliqués dans les cancers du poumon liés au tabac		
Classe chimique	Carcinogènes	Quantité (ng par cigarette)
HAP	Benzo(a)Pyrène	9
	Dibenz(a,h)anthracene	4
Nitrosamines	NNK	123
Hydrocarbures volatiles	Benzène	59 000
	1,3-butadiène	52 000
Autres composés organiques	Oxyde d'éthylène	7 000
Aldéhydes	Formaldéhyde	16 000
	Acétaldéhyde	819 000
Métaux	Cadmium	132

B.



C.

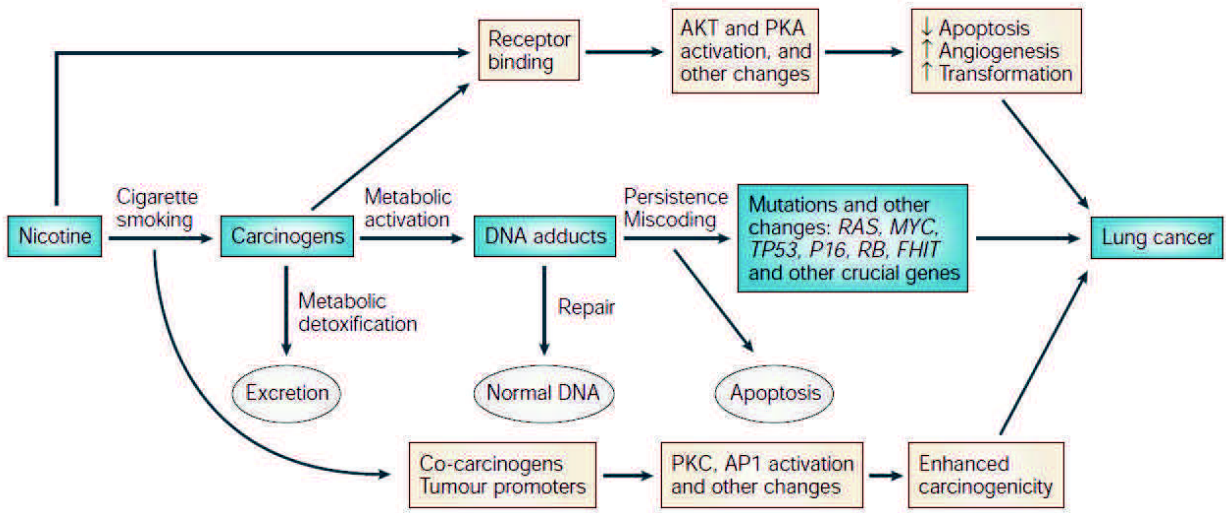


Figure 16 : Carcinogènes du tabac et cancers pulmonaires

A. Principaux carcinogènes de la fumée de cigarette ayant été directement impliqués dans le cancer du poumon.
HAP : Hydrocarbures aromatiques polycycliques ; NKK : Nicotine-derived nitrosamine ketone. *Adapté de (Hecht, 2003).*

B. Exemple de formation d'un adduit à l'ADN : modification covalente d'une guanine par le carcinogène NNK.
POB : pyridyloxobutyl. *Adapté de (Sturla, 2007).*

C. Représentation schématique des voies d'action de la cigarette sur le développement tumoral. *Tiré de (Hecht, 2012).*

Plus récemment, la qualité globale de l'air a été mise en cause ([Chiu et al., 2006](#)) et en juin 2012, l'OMS a fait passer les particules fines issues de moteurs diesel du statut de cancérogène probable à cancérogène certain. Même si les effets sont difficiles à chiffrer et encore discutés, les particules émises par les automobiles seraient responsables de la mort par cancer du poumon de 600 à 1100 personnes par an en France ([travail, 2009](#)).

2.1.2.3. Faible contribution de facteurs génétiques familiaux

Alors que les formes familiales héréditaires de cancer sont relativement courantes et bien étudiées pour certaines localisations tumorales (liens entre rétinoblastome et mutation du gène *RB*, entre cancer du sein et mutation des gènes *BRCA*), elles sont considérées jusqu'à aujourd'hui comme inexistantes pour les cancers du poumon. Néanmoins, des études épidémiologiques récentes ont montré l'existence de « familles à cancers du poumon » ([Gorlova et al., 2007](#)), parmi lesquelles le risque de développer un cancer du poumon serait indépendant des habitudes de vie (tabagisme actif et/ou passif notamment). Certains locus ont également été suggérés comme possiblement liés au cancer du poumon, tels que les locus 6q23-25 ([Bailey-Wilson et al., 2004](#)) ou 15q25, ce dernier comprenant un gène du récepteur nicotinique à acétylcholine ([Hung et al., 2008](#)). Cela suggère l'existence d'une composante héréditaire discrète, non transmise selon les lois de Mendel car interagissant avec l'environnement. Une attention particulière a aussi été portée aux facteurs impliqués dans le métabolisme des carcinogènes du tabac (comme le système de détoxification du cytochrome p450) : certaines mutations transmises entre générations pourraient induire une plus grande sensibilité des individus à l'effet de ces carcinogènes, donc une plus grande susceptibilité au cancer du poumon lié au tabac ([Ariyoshi et al., 2002](#)).

2.2. Carcinomes pulmonaires

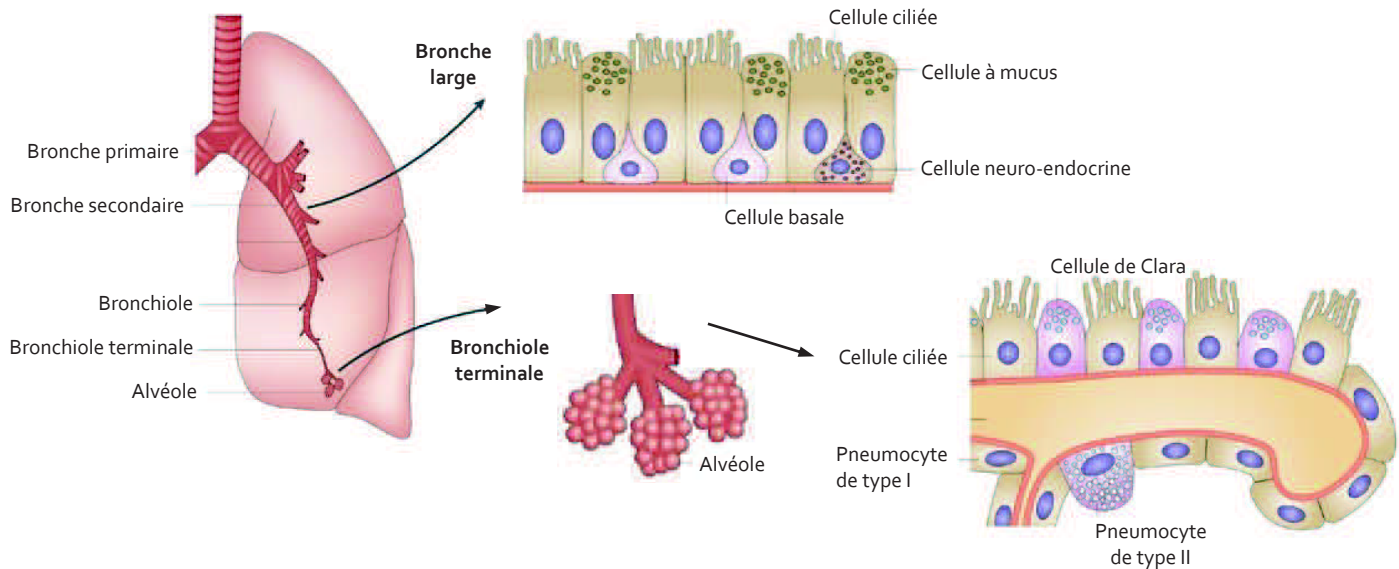
2.2.1. Développement d'un carcinome pulmonaire

Le cancer du poumon est une maladie caractérisée par une prolifération importante et chaotique de cellules anormales dans les tissus du poumon, puis par une invasion d'autres organes par les cellules cancéreuses à l'origine de tumeurs secondaires ou métastases. La grande majorité des cancers du poumon sont des carcinomes, issus de cellules épithéliales pulmonaires. Le carcinome pulmonaire n'est souvent détecté que 5 à 40 ans après l'exposition au carcinogène (**Figure 15B**), suite à un processus de développement tumoral complexe, souvent silencieux et comprenant de multiples étapes.

2.2.1.1. La muqueuse bronchique non pathologique et ses types cellulaires

La muqueuse trachéobronchique est recouverte d'un épithélium pseudostratifié reposant sur une membrane basale (**Figure 17A**). Dans la partie haute des voies respiratoires (trachée, bronches), les cellules à la surface de l'épithélium sont les cellules caliciformes et les cellules ciliées. Les cellules caliciformes (ou cellules à mucus) sont responsables de la plus grande partie de la production de mucus, tandis que le mouvement coordonné des cils des cellules ciliées permet au mucus de remonter jusqu'à la glotte où il est dégluti avec les particules captées. Les cellules basales, ancrées à la membrane basale, assurent par division le renouvellement de l'épithélium de surface, notamment suite à une agression de la muqueuse. Enfin, 4% des cellules de l'épithélium sont des cellules neuroendocrines, capables de sécréter différentes substances (bombésine-like-peptides, calcitonine, sérotonine...). Dans la partie distale des voies respiratoires (bronchioles terminales et alvéoles), l'épithélium s'enrichit des cellules de Clara sécrétrices et de pneumocytes de type I et II. Parmi ces différents types cellulaires, toutes les cellules capables de se diviser sous l'effet d'un processus d'agression chronique ou du fait d'agents carcinogènes (cellules basales, cellules neuroendocrines, cellules de Clara et pneumocytes de type II) peuvent proliférer et devenir néoplasiques ([Brambilla et al., 1996](#)).

A.



B.

CARCINOGENESE

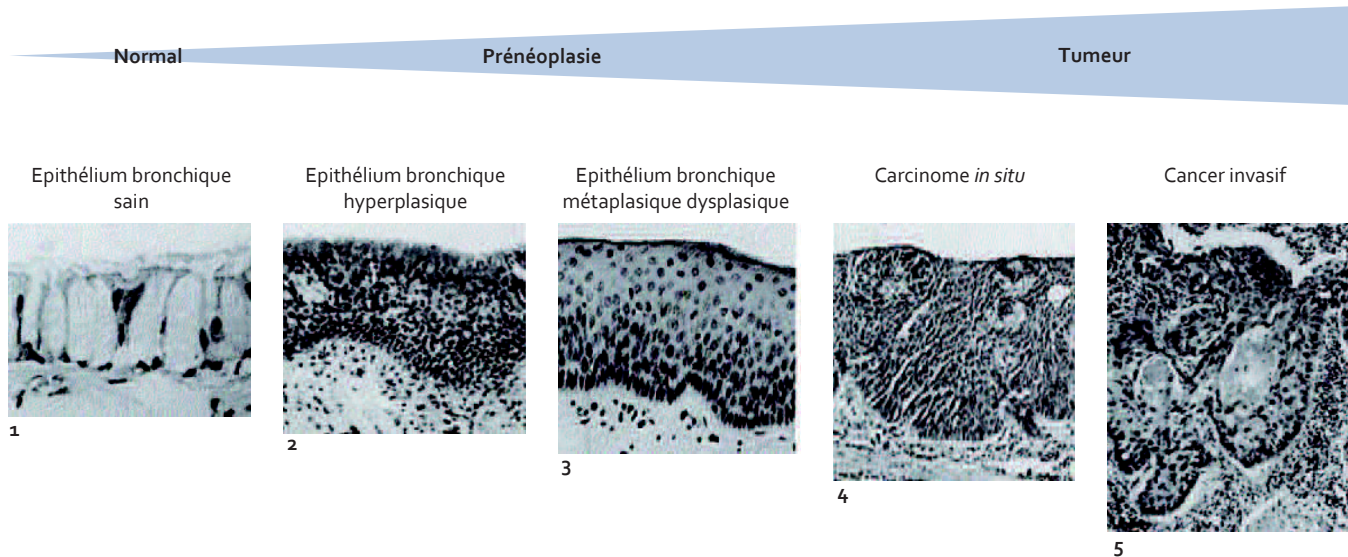


Figure 17 : L'épithélium bronchique normal et pathologique

A. Anatomie simplifiée du poumon et cellules épithéliales pulmonaires à l'origine des tumeurs pulmonaires.

Adapté de (Sun et al., 2007a).

B. Aspects histologiques du développement d'un carcinome bronchique.

Adapté de (Averbach et al., 1961).

1. L'épithélium bronchique est initialement un épithélium pseudostratifié prismatique, fait majoritairement de cellules épithéliales ciliées, de cellules à mucus et de cellules basales, le tout surmontant la lame basale. **2.** L'épithélium s'épaissit par augmentation du nombre de couches cellulaires, ici du fait de la prolifération chaotique des cellules basales. C'est une étape d'hyperplasie épithéliale au cours de laquelle l'épithélium prismatique cilié mucipare pseudostratifié (où toutes les cellules reposent sur la lame basale) se transforme en un épithélium de type épidermoïde pluristratifié. On note la persistance de cellules ciliées et à mucus en surface. **3.** Cette étape de métaplasie est marquée par une désorganisation structurale de l'épithélium: les cellules prismatiques ciliées et à mucus en surface disparaissent au profit d'un épithélium pavimenteux constitué de cellules d'origine basale. Puis la métaplasie devient dysplasique: les cellules basales qui prolifèrent de manière aberrante accumulent des atypies cellulaires et les défauts nucléaires. **4.** Le carcinome est dit *in situ* car les cellules cancéreuses restent entourées par la lame basale. **5.** Enfin, la lame basale est percée. Les cellules cancéreuses envahissent le tissu conjonctif.

2.2.1.2. Aspects histologiques de la carcinogenèse pulmonaire

Les lésions précancéreuses de la muqueuse bronchique sont dans la plupart des cas le résultat d'une agression chronique de l'épithélium par les cancérogènes de la fumée de cigarette. Les étapes de développement d'un carcinome bronchique lié au tabac ont depuis longtemps été décrites sur le plan histologique ([Auerbach et al., 1961](#) ; [Foulds, 1954](#)). Le carcinome se développe au sein de l'épithélium bronchique via une séquence de changements morphologiquement reconnaissables (**Figure 17B**). L'épithélium, initialement pseudostratifié, s'épaissit lors de la phase d'hyperplasie épithéliale du fait de la prolifération chaotique des cellules basales. Au stade métaplasique, l'épithélium devenu pluristratifié se désorganise, et les cellules ciliées et les cellules à mucus en surface disparaissent. Puis la métaplasie devient dysplasique lorsque les cellules basales qui prolifèrent accumulent des atypies cellulaires. Les lésions observées jusqu'ici sont précancéreuses : de petite taille (1 à 3 millimètres), strictement limitées à l'épithélium, elles peuvent même se résorber si la source d'irritation chronique (tabac) est supprimée. A un stade plus tardif, les cellules basales ont acquis un phénotype cancéreux plus marqué (prolifération non limitée, perte de la mort cellulaire). Le carcinome est dit *in situ* car les cellules cancéreuses restent entourées par la lame basale ; cette tumeur localisée peut être supprimée par chirurgie. Enfin, la lame basale est percée. Les cellules cancéreuses envahissent le tissu conjonctif, rejoignent les vaisseaux sanguins et lymphatiques et peuvent proliférer au sein d'un autre organe de l'organisme en formant des métastases, préférentiellement vers le foie (33 à 40% des cas), les glandes surrénales (18 à 38%), le cerveau (15 à 43%), les os (19 à 33%) et la moëlle osseuse. Cette séquence de transformations successives des lésions depuis l'hyperplasie jusqu'au carcinome invasif est un modèle de transformation horizontale, avec une évolutivité variable et une possibilité d'arrêt à chaque stade.

2.2.1.3. La stroma-réaction

Les tumeurs pulmonaires ne sont pas composées uniquement de cellules malignes. A un stade avancé du développement tumoral, les cellules cancéreuses sont entourées par un tissu de soutien non tumoral appelé stroma. Celui-ci est composé d'une matrice extracellulaire et d'un compartiment cellulaire comprenant des fibroblastes, des cellules endothéliales et sanguines, des lymphocytes T (Tumor Infiltrating Lymphocytes TIL) et des macrophages (Tumor Associated Macrophages TAM). Pour les cellules cancéreuses, le stroma assure d'abord un support mécanique et structural : les fibroblastes synthétisent les composantes de la matrice extracellulaire qui constitue un réseau de protéines fibreuses (collagène, élastine, fibronectine) et de polysaccharides. Les interactions physiques, notamment entre fibronectine de la matrice et intégrines, pourraient stimuler la croissance cellulaire des cellules cancéreuses ([Ritzenthaler et al., 2008](#)). Ce microenvironnement tumoral contribue également au développement de la tumeur en favorisant son approvisionnement en nutriments et sa dispersion par angiogenèse : la formation de nouveaux vaisseaux sanguins est permise par la sécrétion de médiateurs pro-angiogéniques (protéines de la famille du vascular endothelial growth factor vEGF principalement) induisant la prolifération des cellules endothéliales ([de Mello et al., 2012](#)). Enfin, les cellules du microenvironnement modulent la différenciation, la prolifération, l'adhésion et la capacité migratoire des cellules cancéreuses, via la sécrétion de nombreux facteurs solubles (facteurs de croissance, cytokines, chimiokines) ([Bremnes et al., 2011](#)). Ces interactions tumeur-stroma se révèlent ainsi être de nouvelles cibles thérapeutiques dans le traitement des tumeurs pulmonaires ([El-Nikhely et al., 2012](#)).

2.2.2. Classification des cancers du poumon

La classification CIM-O-3 (Classification internationale des maladies oncologiques de l'OMS) ([Fritz et al., 2008](#)) propose deux classifications distinctes des tumeurs pulmonaires : une classification histopathologique basée sur les caractéristiques morphologiques de la tumeur et des cellules cancéreuses, doublée d'une classification selon le stade qui décrit l'extension de la pathologie néoplasique. Ces deux

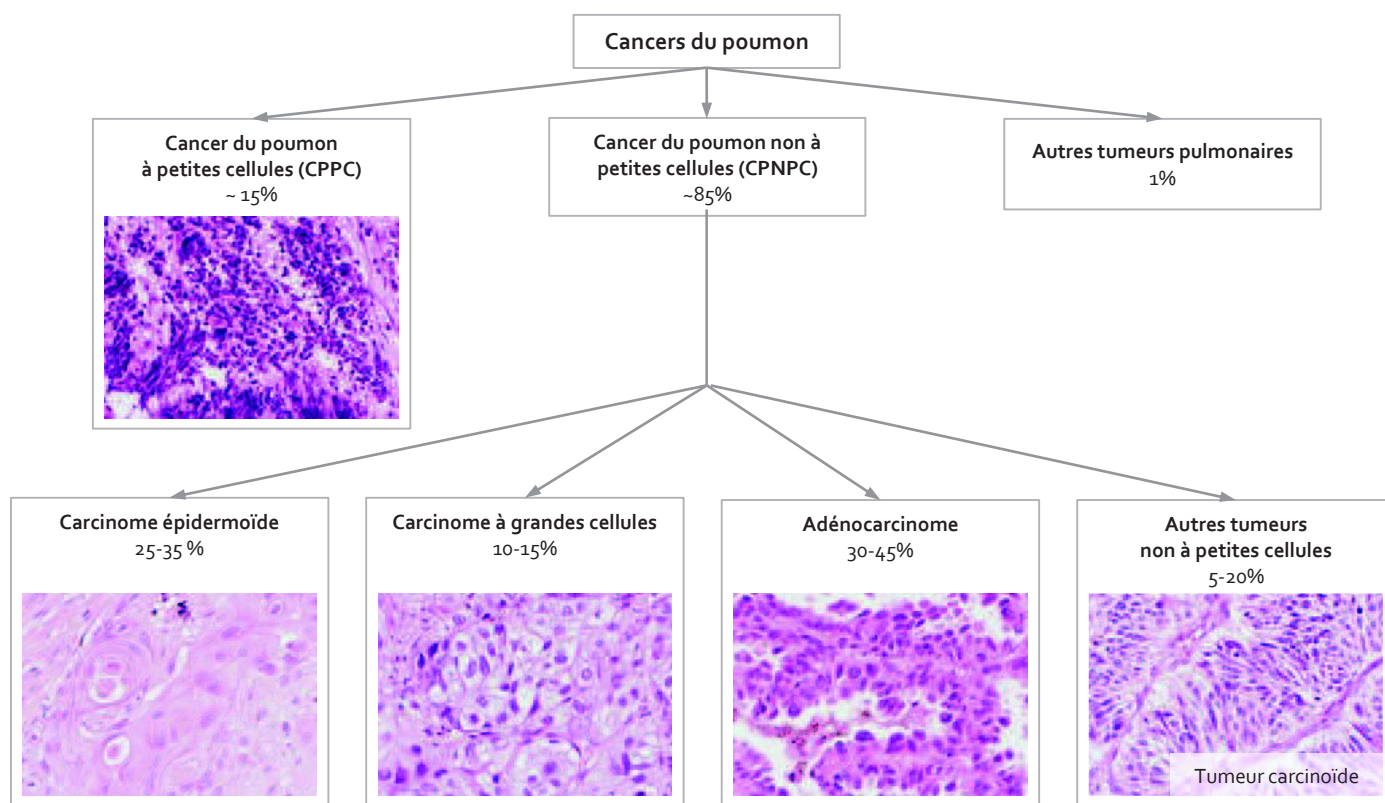


Figure 18 : Classification histopathologique simplifiée des tumeurs pulmonaires

Classification OMS de 2004.

Adapté de (Travis et al., 2004).

Photographies tirées de (Motadi et al., 2007).

classifications prennent en compte l'hétérogénéité des tumeurs pulmonaires et guident la prise en charge clinique des patients.

2.2.2.1. Classification histopathologique des tumeurs pulmonaires

D'un point de vue histopathologique, le cancer du poumon est un néoplasme très complexe. Contrairement à de nombreux autres organes, les poumons sont le siège d'un champ très large de tumeurs épithéliales qui varient en terme de localisation (trachée, bronche, bronchiole, alvéole) et d'histologie. Néanmoins, on distingue deux types principaux de carcinomes du poumon : les carcinomes non à petites cellules et les carcinomes à petites cellules (**Figure 19**). Ces deux classes sont ensuite subdivisées en de nombreux sous-types sur des critères histopathologiques (taille et aspect des cellules malignes), selon la classification simplifiée de l'OMS ([Travis et al., 2004](#)). Cette classification est fondamentale d'un point de vue clinique : elle a des implications importantes sur la prise en charge, le pronostic et le traitement de la maladie ([Vaporciyan et al., 2002](#)). Et au-delà du caractère initial uniquement histologique, on découvre aujourd'hui que cette classification revêt une vraie pertinence biologique et génétique : les différents sous-types présentent des caractéristiques moléculaires distinctes (**voir paragraphe 2.3.3 et Figure 31**) et peuvent correspondre à des origines cellulaires différentes (**voir paragraphe 2.2.2.1.3**).

2.2.2.1.1. Les carcinomes bronchiques non à petites cellules CBNPC

Les carcinomes non à petites cellules représentent plus de 80% des cas de cancer du poumon. Ils constituent un groupe hétérogène, au sein duquel on distingue 3 sous-types principaux (**Figure 19**).

Les carcinomes épidermoïdes bronchiques CEB (*squamous-cell lung carcinoma SQC-LC*), qui représentent 30% des CBNPC, sont très fortement associés au tabagisme. Ils se développent généralement dans la lumière d'une bronche majeure sous la forme d'une masse tumorale végétante. Les CEB sont bien différenciés et les cellules cancéreuses présentent les caractéristiques de l'épithélium trachéobronchique (kératinisation, ponts intercellulaires, cytoplasme abondant avec réticulum endoplasmique et appareil de Golgi développés). Ils sont divisés en quatre sous-types ([Travis et al., 2004](#)) : papillaire, basaloïde, à cellules claires, à petites cellules.

Les adénocarcinomes (ADC – 40% des CBNPC), qui débutent généralement dans des régions situées plus à l'extrémité de l'arbre bronchique (bronchiole terminale et alvéole), peuvent exhiber une différenciation glandulaire de degré variable. Ils sont divisés en plusieurs sous-types ([Travis et al., 2004](#)) (papillaire, solide, acinaire, bronchioalvéolaire...) mais la plupart des ADC sont hétérogènes, avec présence d'au moins deux sous-types histologiques. Les femmes développent davantage d'adénocarcinomes que les hommes ([Patel, 2005](#)) et l'ADC est la forme la plus fréquente de cancer du poumon chez les personnes n'ayant jamais fumé ([Subramanian and Govindan, 2007](#)).

Enfin, les carcinomes pulmonaires à grandes cellules CPGC (*large-cell lung carcinoma LC-LC* - 9%) sont ainsi nommés car ils sont formés de cellules cancéreuses de grande taille, à cytoplasme abondant et à gros noyau avec nucléole proéminent. Les sous-types de CPGC comprennent notamment les carcinomes basaloïdes, les carcinomes à cellules claires ou les carcinomes neuroendocrines à grandes cellules ([Travis et al., 2004](#)). A ces trois sous-types principaux de CBNPC s'ajoutent d'autres sous-types plus marginaux comme les tumeurs carcinoïdes, les carcinomes adénoquameux et sarcomatoïdes ou encore les tumeurs des glandes bronchiques ([Travis et al., 2004](#)). La fréquence relative des différents types histologiques de CBNPC a évolué avec le temps. Chez les hommes, les carcinomes épidermoïdes tendent depuis 25 ans à diminuer au profit des adénocarcinomes. Chez les femmes, tous les types histologiques progressent en terme d'incidence, mais la prédominance des adénocarcinomes tend à s'accroître avec le temps ([Devesa et al., 2005](#)).

A. T : Taille de la tumeur primitive

TX	Tumeur qui ne peut être évaluée . Tumeur juste démontrée (présence de cellules malignes dans les expectorations et sécrétions bronchiques) mais non visualisée par des examens endoscopiques ou d'imagerie .
To	Pas d'évidence de tumeur primitive.
Tis	Carcinome <i>in situ</i>
T1	Tumeur inférieure u égale à 3 cm dans sa plus grande dimension, entourée par le poumon ou la plèvre viscérale, sans évidence d'invasion de la bronche souche.
T2	Tumeur de plus de 3 cm sans dépasser 7 cm dans sa plus grande dimension. Tumeur présentant une des caractéristiques suivantes : atteinte de la bronche souche à 2 cm ou plus de la carène, ou invasion de la plèvre viscérale, ou présence d'une atélectasie ou d'une pneumopathie obstructive s'étendant à la région hilare sans atteindre le poumon complet T2a : Tumeur de plus de 3 cm sans dépasser 5 cm ; T2b : Tumeur de plus de 5 cm sans dépasser 7 cm
T3	Tumeur de plus de 7 cm . Tumeur envahissant directement une des structures suivantes : la paroi thoracique , le diaphragme, le nerf phrénique, la plèvre médiastinale ou le péricarde pariétal. Tumeur dans la bronche souche à moins de 2 cm de la carène sans l'envahir, ou associée à une atélectasie ou une pneumopathie obstructive du poumon entier . Tumeur avec présence d'un nodule tumoral distinct dans le même lobe.
T4	Tumeur de toute taille envahissant directement une des structures suivantes : médiastin, cœur, gros vaisseaux, trachée, nerf laryngé récurrent, œsophage, corps vertébral, carène . Tumeur avec présence d'un nodule tumoral distinct dans un autre lobe du poumon atteint.

B. N : Adénopathies régionales

NX	Les ganglions régionaux ne peuvent être évalués
No	Pas de métastase ganglionnaire lymphatique régionale
N1	Métastase(s) dans les ganglions lymphatiques intrapulmonaires, péribronchiques et/ou hilaires homolatéraux
N2	Métastase (s) dans les ganglions lymphatiques médiastinaux homolatéraux et/ou sous carinaires
N3	Métastase(s) dans les ganglions lymphatiques médiastinaux controlatéraux, hilaires controlatéraux, scalènes ou sous-claviculaires

C. M : Métastases à distance

MX	La présence de métastase(s) à distance ne peut être évaluée
Mo	Absence de métastase à distance
M1	Présence de métastase(s) à distance

Figure 19 : Grille de classification TNM des tumeurs pulmonaires (Classification 2009, 7^{ème} édition)
Tiré de (Goldstraw et al., 2007).

2.2.2.1.2. Les carcinomes bronchiques à petites cellules CBPC

Les carcinomes à petites cellules représentent environ 15% des cas de cancer du poumon, soit 5000 nouveaux cas par an en France. Ce type de cancer est fortement associé à la cigarette : il touche essentiellement des personnes d'âge moyen (50-60 ans) dont 95% avec des antécédents tabagiques. Le CBPC est la plus agressive des tumeurs pulmonaires, avec un taux de survie à 5 ans compris entre 1 et 5% et une survie moyenne d'environ 6 à 10 mois après le diagnostic (Salcido et al., 2010). Les CBPC grossissent rapidement (Rosti et al., 2006), ont un plus grand potentiel métastatique que les CBNPC (Varker, 2003) et en sont souvent à un stade invasif au moment de la détection (60 à 70% de tumeurs métastatiques s'étendant au-delà du thorax) (Rosti et al., 2006). Les CBPC ont tendance à survenir au niveau central du poumon, dans les voies aériennes les plus grandes (bronches primaires et secondaires). Comme leur nom l'indique, les cellules de CBPC sont des cellules de petite taille, à cytoplasme étroit voire quasiment absent. Les cellules de CBPC exhibent couramment un caractère neuroendocrine (Linnoila, 1996).

2.2.2.1.3. Des origines cellulaires différentes ?

Déterminer le type cellulaire à l'origine d'un cancer est une tâche complexe, particulièrement dans le cas des cancers du poumon qui sont le plus souvent diagnostiqués à des stades avancés (Minna et al., 2002). Pour la plupart des types de carcinomes pulmonaires, les cellules cancéreuses présentent des caractéristiques proches de celles des cellules différenciées de la région où la tumeur s'est développée (Figure 17A). Par exemple, les carcinomes épidermoïdes présentent la kératinisation caractéristique des cellules épithéliales de la trachée et des voies aériennes proximales, les adénocarcinomes présentent des marqueurs cellulaires caractéristiques des cellules de Clara ou des pneumocytes de type II (Kim et al., 2005), tandis que les CBPC présentent souvent des caractéristiques proches des cellules neuroendocrines, dispersées à l'état normal dans tout l'épithélium bronchique (Dasanu et al., 2010). Toutefois, la vision selon laquelle les différents types histologiques de cancers pulmonaires dérivent de lignages cellulaires distincts et en conservent indéfiniment les propriétés est progressivement remise en cause : les caractéristiques des cellules cancéreuses ne tiendraient pas tant d'un type cellulaire originel que de l'influence du microenvironnement tumoral (Kim et al., 2005). On pense aujourd'hui que les ADC ne dérivent pas de la transformation maligne de cellules de Clara ou de pneumocytes de type II, mais plutôt de cellules souches bronchioalvéolaires, responsables du maintien des cellules de Clara et des pneumocytes de type II (Kim et al., 2005). D'autre part, on observe parfois lors de la progression de CBPC une transition vers un phénotype CBNPC, notamment après une rechute, lorsque les cellules résistantes du CBPC présentent les caractéristiques morphologiques typiques des carcinomes à grandes cellules (Mabry et al., 1991).

2.2.2.2. Classification selon le stade

Pour préciser le diagnostic et affiner la prise en charge thérapeutique d'une tumeur pulmonaire, les cliniciens ajoutent à la classification histopathologique une classification selon le stade (Figure 21), elle-même basée sur la classification TNM (Figure 19). La classification TNM est un système international, proposé par le chirurgien français Pierre Denoix, de classement des cancers selon leur extension anatomique. La 7^{ème} édition révisée a été publiée en 2009 (Goldstraw et al., 2007). Le premier paramètre (T) décrit la taille de la tumeur primitive : ce paramètre va d'une tumeur primitive non détectée (T₁) à une tumeur envahissante (T₄) (Figure 19A). Le second paramètre (N) décrit l'envahissement des ganglions lymphatiques régionaux par les cellules cancéreuses : il va d'une absence de métastase ganglionnaire (N₀) à un envahissement des ganglions contralatéraux et lointains (N₃) (Figure 19B). Le dernier paramètre (M) décrit de manière binaire l'invasivité du cancer : pas de métastase (M₀) ou présence de métastases à distance (M₁) (Figure 19C). On utilise alors de manière combinatoire ces trois paramètres TNM pour établir le stade (Figure 20). Le cancer évolue depuis le stade IA (le meilleur pronostic), IB ou II jusqu'au stade IV (le cancer métastatique ou généralisé de pronostic très sombre) (Mountain et al., 2003).

Cancer occulte	TX	No	Mo
Stade o	Tis	No	Mo
Stade IA	T1	No	Mo
Stade IB	T2	No	Mo
Stade IIA	T1	N1	Mo
	T2a	N1	Mo
	T2b	No	Mo
Stade IIB	T2b	N1	Mo
	T3	No	Mo
Stade IIIA	T1, T2	N2	Mo
	T3	N1, N2	Mo
	T4	No, N1	Mo
Stade IIIB	T4	N2	Mo
	Tout T	N3	Mo
Stade IV	Tout T	Tout N	M1

Figure 20 : Grille de classification par stades des tumeurs pulmonaires (Classification 2009, 7^{ème} édition)
Tiré de (Goldstraw et al., 2007).

2.3. Carcinogénèse pulmonaire

2.3.1. Altérations de facteurs cellulaires clés : oncogènes et suppresseurs de tumeur

L'exposition à des carcinogènes environnementaux constitue l'évènement fondateur le plus courant de la pathogénèse pulmonaire (Hecht, 2012). Il est maintenant clairement démontré qu'une exposition chronique aux carcinogènes de la fumée de cigarette entraîne la formation d'adduits à l'ADN, source d'une accumulation au cours du temps d'altérations dans le génome des cellules bronchiques (voir paragraphe 2.1.2.1). En réponse à l'action d'un carcinogène, les cellules de l'épithélium bronchique possèdent différents mécanismes de protection tels que les systèmes de détoxification et d'excrétion de la substance (Luch and Baird, 2005), les systèmes cellulaires de réparation des dommages à l'ADN (Wei et al., 2000) ou encore l'élimination des cellules anormales par apoptose. Tous ces phénomènes permettent en théorie de limiter les dégâts subis et donc de prévenir la transformation maligne. Mais si les dommages à l'ADN persistent et que la cellule réplique un ADN endommagé, il y a risque de changements définitifs dans la séquence d'ADN et altération irréversible du génome. Certaines de ces mutations peuvent conférer un avantage à la cellule cancéreuse qui les porte, comme une capacité de division accrue ou une plus grande résistance aux signaux de mort cellulaire : ces mutations seront donc sélectionnées au cours du développement tumoral, la cellule se divisant activement jusqu'à constituer un clone. C'est en particulier le cas des mutations se produisant dans la séquence de gènes cruciaux pour le contrôle de la croissance cellulaire, comme les oncogènes ou les gènes suppresseurs de tumeurs (Hecht, 2012). Dans cette première partie introductive, nous ne présenterons comme preuve de concept que deux gènes mutés dans les cancers pulmonaires, le proto-oncogène *KRAS* et le suppresseur de tumeur *TP53*. Les altérations dans l'expression de nombreux autres gènes-clés seront abordées plus en détails dans la troisième partie de cette introduction (voir paragraphe 3.2).

2.3.1.1. Activations d'oncogènes

Le terme « oncogène » est apparu en 1969 (Huebner and Todaro, 1969) suite à la découverte du rôle de *src*, le premier oncogène identifié chez le rétrovirus RSV du poulet (Martin, 1970). La découverte du rôle fondamental des oncogènes a été récompensée par le Prix Nobel de Médecine et Physiologie en 1989 (Bishop et Varmus). Aujourd'hui, plus de 40 proto-oncogènes humains ont été décrits et certains d'entre eux jouent un rôle majeur dans la carcinogénèse bronchique. Lorsqu'une mutation activatrice se produit dans la séquence nucléotidique d'un proto-oncogène, la version mutée du gène est alors appelée oncogène. Ce phénomène peut conduire à la production d'une protéine mutée hyperactive ou constitutivement active (appelée oncoprotéine). Le plus souvent, ces protéines sont impliquées dans des voies stimulant la division cellulaire et ont pour fonction d'accélérer la prolifération ; la protéine mutée conduit donc à l'activation continue des voies de prolifération, sans que la cellule ne nécessite les signaux qui stimulent normalement la croissance cellulaire. L'évènement génétique conduisant à l'activation d'un oncogène peut ne se produire que sur un seul allèle du proto-oncogène, étant donné que la plupart des mutations activatrices sont dominantes (Figure 21A).

KRAS est l'un des oncogènes les plus couramment activés dans les cancers du poumon (Sato et al., 2007). De 10 à 25% des CBNPC, et jusqu'à 30% des adénocarcinomes bronchiques, présentent une mutation ponctuelle du proto-oncogène *KRAS* (Brose et al., 2002 ; Riely et al., 2008). Chez les fumeurs, ces mutations sont majoritairement de type transversions G vers T (Figure 21B), une mutation caractéristique de l'exposition aux carcinogènes du tabac, induite par le benzo(a)pyrène *in vitro* (Hu et al., 2003). Ces mutations ponctuelles touchent plus particulièrement le codon 12 de la séquence génique de *KRAS* (le codon sauvage GGT devient GTT), conduisant à la production d'une protéine K-ras mutée (une valine est incorporée à la place d'une glycine) (Keohavong et al., 1996). On a cherché à comprendre l'implication biologique de cette substitution. K-ras est une petite protéine G associée à la membrane, qui régule les voies de transduction impliquées dans la prolifération et la différenciation. Les mutations ponctuelles au niveau du codon 12 peuvent conduire à la

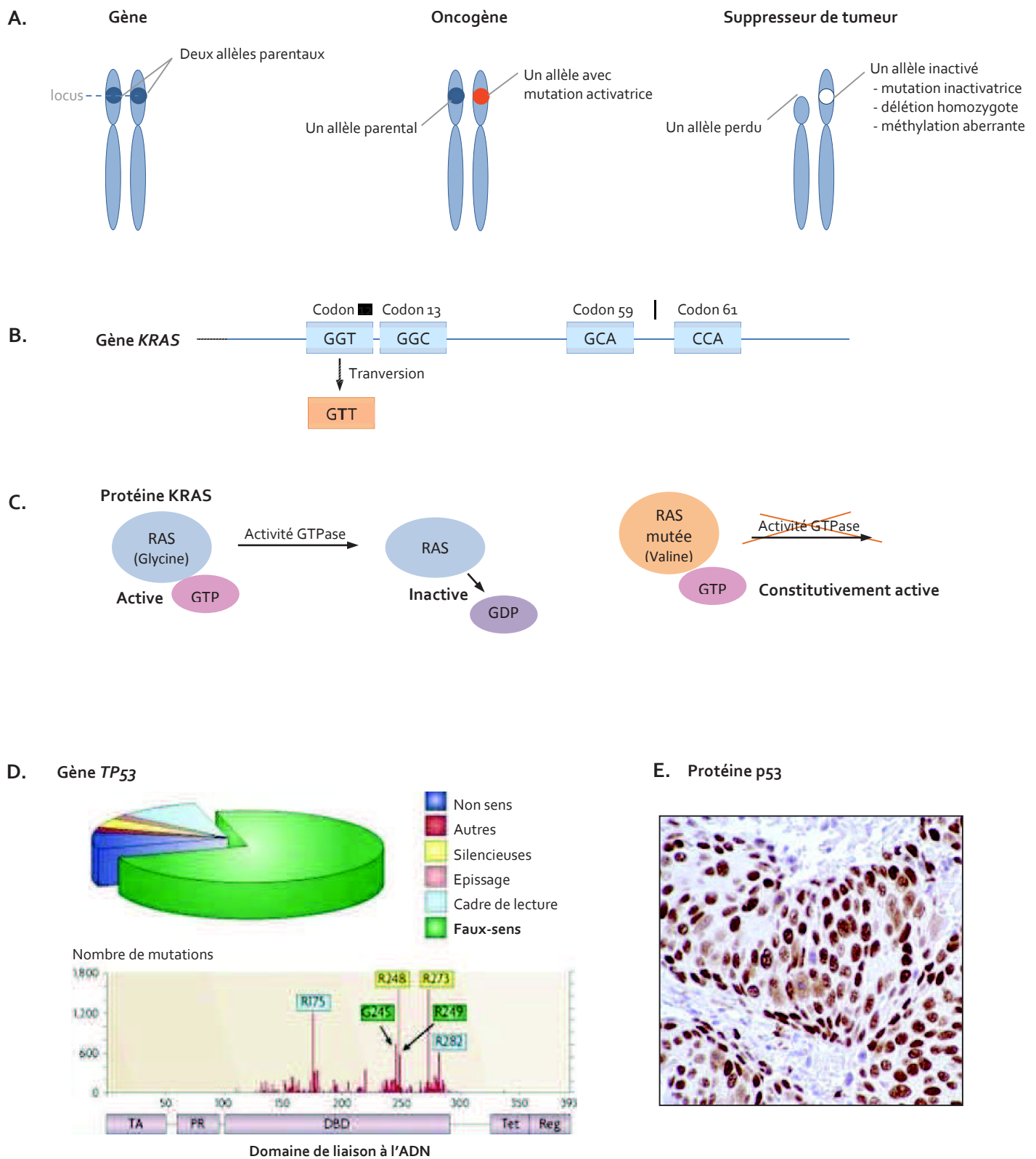


Figure 21 : Mutations de l'oncogène KRAS et du suppresseur de tumeur p53 dans les cancers pulmonaires

A. Principe d'activation d'un oncogène et d'inactivation d'un suppresseur de tumeur

B. Mutations activatrices de type transversions du gène KRAS : les codons 12, 13, 59 et 61, correspondant aux sites de fixation de GTP/GDP sur la protéine, sont fréquemment mutés.

C. La protéine KRAS est une protéine G cytoplasmique active sous sa forme liée à la guanosine-triphosphate (GTP). La protéine KRAS mutée voit sa conformation modifiée et son activité GTPasique diminuée. Plus capable d'hydrolyser le GTP et de se dissocier du GDP, elle est constitutivement active et transmet un signal prolifératif.

D. Distribution des mutations ponctuelles inactivatrices du gène TP53 . Il s'agit principalement de mutations faux-sens entre les codons 120 et 300, dans la région codant le domaine de liaison à l'ADN (DNA-binding domain DBD).

Traduit de (Brosh and Rotter, 2009).

E. Accumulation nucléaire de la protéine p53 mutée détectée par immunohistochimie sur un carcinome épidermoïde.

Tiré de (Lantuéjoul, 2006).

production d'une protéine K-ras mutée qui n'est plus capable de se dissocier du GTP et qui se maintient donc dans une conformation activée (Rodenhuis and Slebos, 1992) (Figure 21C). Elle active alors de manière constitutive la voie de croissance cellulaire RAS/RAF/MEK/ERK en aval (Figure 23) (Ihle et al., 2012). Les mutations de KRAS sont associées à un pronostic sombre et une faible survie des patients concernés (Husgafvel-Pursiainen et al., 1993 ; Lim et al., 2009).

A cette définition *stricto sensu* d'un oncogène, on adjoint également les gènes codant des protéines promouvant la transformation qui ne sont pas mutés, mais surexprimés ou exprimés de manière aberrante, notamment suite à des événements d'amplification génique ou de translocation, ou du fait de perturbations épigénétiques. Ces aspects seront abordés dans la deuxième partie de cette introduction qui traitera plus particulièrement des dérégulations des expressions géniques dans les cancers pulmonaires (voir paragraphe 3.2.2).

2.3.1.2. Inactivations génétiques de gènes suppresseurs de tumeur

Les gènes suppresseurs de tumeurs, à l'opposé des proto-oncogènes, codent des protéines réprimant la prolifération cellulaire et promouvant l'apoptose. Lorsque les fonctions des suppresseurs de tumeur ne sont pas assurées, la cellule cancéreuse devient insensible aux signaux anti-prolifératifs et inducteurs d'apoptose, et se divise de manière incontrôlée.

Le rôle essentiel de « gardien du génome » de la phosphoprotéine nucléaire p53 se manifeste notamment après une hypoxie ou un dommage subi par l'ADN. Tout d'abord, p53 assure le blocage du cycle cellulaire en G₁ : p53 induit l'expression de p21/WAF1, un inhibiteur des complexes Cycline D1/CDK4 qui limite la phosphorylation de Rb et donc l'entrée en phase S (Harper et al., 1993). Ce blocage en G₁ permet aux processus de réparation de l'ADN d'intervenir, mais si les dommages sont trop importants ou non réparables, p53 est capable d'induire l'apoptose en modifiant le niveau transcriptionnel des gènes BCL2 et BAX (Reed et al., 1996) (Figure 23). La survenue d'une anomalie de fonctionnement de p53 est alors susceptible de conduire à une dérégulation des voies de prolifération et des voies d'apoptose, aboutissant au passage des cellules de G₁ en S et au blocage de la mort cellulaire. Les altérations du gène *TP53* sont les anomalies génétiques les plus fréquemment retrouvées dans les cancers humains (50 à 60% des cas toutes localisations tumorales confondues), dont les cancers du poumon (40 à 70% des cas de cancer du poumon, dont environ 90% des cas de CBPC) (Bodner et al., 1992 ; Takahashi et al., 1989), plus fréquemment chez les fumeurs (Pfeifer et al., 2002 ; Ryberg et al., 1994).

Dans la majorité des cas où le gène *TP53* n'est pas fonctionnel, on observe une délétion allélique au niveau du locus 17p13 ; puis c'est à l'occasion de la survenue d'une mutation ponctuelle sur le deuxième allèle que la fonction suppressive de p53 est perdue. Il s'agit le plus souvent de mutations ponctuelles faux-sens (transversions G-T ou G-C) entre les positions 120 et 300 (Hollstein et al., 1991) (Figure 21D). La plupart d'entre elles induisent un changement de conformation de la protéine dans le domaine de liaison à l'ADN et supprime ainsi la faculté de la protéine à se lier à ses séquences spécifiques sur le génome (Kim and Deppert, 2006). De manière intéressante, il a été démontré que la formation d'adduits de benzo-(a)-pyrène au niveau des guanines de l'ADN se fait particulièrement au niveau de certains de ces codons « points chauds » mutationnels (positions 157, 248 et 273), induisant des erreurs répliquatives de type transversion (Denissenko et al., 1996) : ainsi, les mutations de *TP53* seraient majoritairement dues au rôle mutagène des composants de la fumée de cigarette (Hainaut and Pfeifer, 2001). La protéine p53 mutée possède parfois une stabilité accrue : sa demi-vie plus longue permet l'accumulation de la protéine non-fonctionnelle dans les cellules cancéreuses, rendant son immunodétection aisée (Figure 21E) (Brambilla et al., 1993).

Il faut noter que les deux allèles d'un gène suppresseur de tumeur doivent être altérés pour que la fonction de la protéine suppresseur de tumeur soit perdue. L'un des allèles peut être transmis par un parent sous une forme mutée, ce qui fait de certains gènes suppresseurs de tumeurs des gènes de susceptibilité

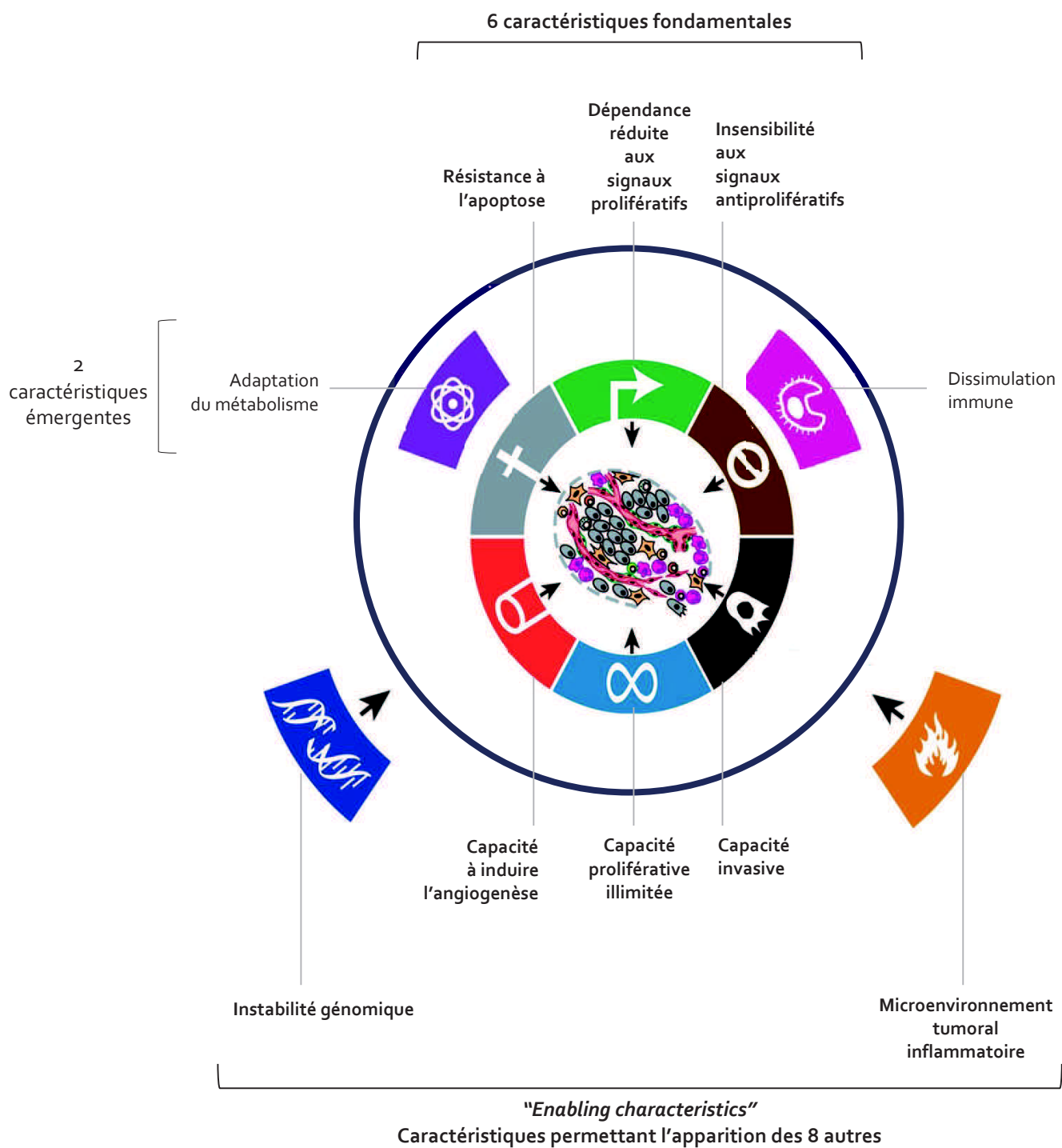


Figure 22 : Les caractéristiques fondamentales des cellules cancéreuses selon Hanahan et Weinberg
Traduit et adapté de (Hanahan and Weinberg, 2000) et (Hanahan and Weinberg, 2011).

associés à certains cancers familiaux. C'est le cas dans les cancers du sein pour les gènes *BRCA*, dont les produits sont impliqués dans la réparation de l'ADN. Mais ce phénomène semble marginal dans les cancers bronchiques (**voir paragraphe 2.1.2.3**) où les altérations des allèles sont acquises au cours de la transformation cellulaire. Cependant, les processus d'inactivation sont souvent différents d'un allèle à l'autre : dans le cas de *TP53*, la perte du premier allèle par délétion (phénomène que l'on nomme « perte d'hétérozygotie » ou "loss of heterozygosity" LOH) est souvent suivie par l'apparition d'une mutation ponctuelle sur le second allèle (**Figure 21A**). Mais au-delà de ces altérations génétiques, on a pris conscience plus récemment de l'importance des phénomènes épigénétiques (tels que les hyperméthylations des promoteurs) dans l'inactivation des gènes suppresseurs de tumeur ; ce phénomène sera discuté dans la troisième partie de cette introduction (**voir paragraphe 3.2.1**).

2.3.2. Détournement de fonctions cellulaires clés et voies cellulaires impliquées dans la carcinogenèse pulmonaire

Suite à la découverte des oncogènes et suppresseurs de tumeur et grâce aux nouvelles méthodes d'analyse génétique des tumeurs pulmonaires, on a progressivement élucidé les mécanismes moléculaires et les processus biologiques conduisant à l'acquisition du phénotype cancéreux par les cellules bronchiques. Le comportement anormal des cellules cancéreuses pulmonaires, maintenant bien décrit à l'échelle phénotypique, se caractérise par des abolitions ou détournements de fonctions cellulaires clés – et il apparaît clair que ce phénotype cancéreux découle directement des altérations moléculaires subies par les cellules bronchiques ([Sato et al., 2007](#)). Nous allons tenter ici de donner un aperçu des caractéristiques fondamentales acquises par les cellules épithéliales au cours de la carcinogenèse pulmonaire, en lien avec quelques voies de signalisation cellulaires fréquemment altérées dans les cancers bronchiques.

Malgré la complexité des pathologies tumorales et l'abondance de la littérature, Hanahan et Weinberg ont suggéré que le comportement cellulaire cancéreux pouvait être décrit de manière simple et que la transformation cellulaire résulterait de l'acquisition par les cellules cancéreuses de six caractéristiques fondamentales ([Hanahan and Weinberg, 2000](#)) (**Figure 22**) : l'affranchissement envers les signaux prolifératifs, l'insensibilité aux signaux anti-prolifératifs, la capacité proliférative illimitée, la résistance à l'apoptose, la capacité à induire l'angiogenèse (permettant d'assurer l'approvisionnement de la tumeur en oxygène par de nouveaux vaisseaux), et l'acquisition d'une capacité invasive (autorisant la dissémination des cellules cancéreuses dans l'organisme et la formation de tumeurs secondaires). A ces six caractéristiques fondamentales énoncées en 2000, les auteurs ont ajouté récemment deux caractéristiques émergentes : la capacité de reprogrammer le métabolisme cellulaire pour supporter efficacement la prolifération néoplasique, ainsi que la capacité des cellules cancéreuses à se dissimuler de la surveillance du système immunitaire assurée par les lymphocytes T, B et NK et les macrophages ([Hanahan and Weinberg, 2011](#)). Selon les auteurs, toutes ces caractéristiques fondamentales doivent être acquises pour assurer la croissance tumorale. Elles apparaîtraient de manière séquentielle et seraient sélectionnées selon un mode analogue à la sélection darwinienne : l'acquisition d'une de ces caractéristiques confère un avantage, qui permet à la cellule de proliférer, formant un clone au sein duquel apparaît une nouvelle capacité, elle aussi sélectionnée, le tout conduisant à la transition progressive d'une cellule saine vers une cellule maligne agressive. L'acquisition de chacune de ces caractéristiques est soutenue par deux mécanismes appelés « enabling characteristics » : la contribution du microenvironnement tumoral d'une part (**voir paragraphe 2.2.1.3**), et l'instabilité génétique d'autre part (**voir paragraphe 3.2.1**) ([Hanahan and Weinberg, 2011](#)). Pour aller plus loin que ces aspects descriptifs, les efforts de recherche se sont orientés vers la compréhension des bases moléculaires conduisant à l'acquisition du phénotype anormal des cellules cancéreuses. Nous allons nous intéresser plus particulièrement à deux caractéristiques des cellules cancéreuses (la prolifération accrue et la résistance à l'apoptose) en décrivant quelques altérations de voies de signalisation cellulaires fréquentes dans les cancers pulmonaires aboutissant à leur acquisition.

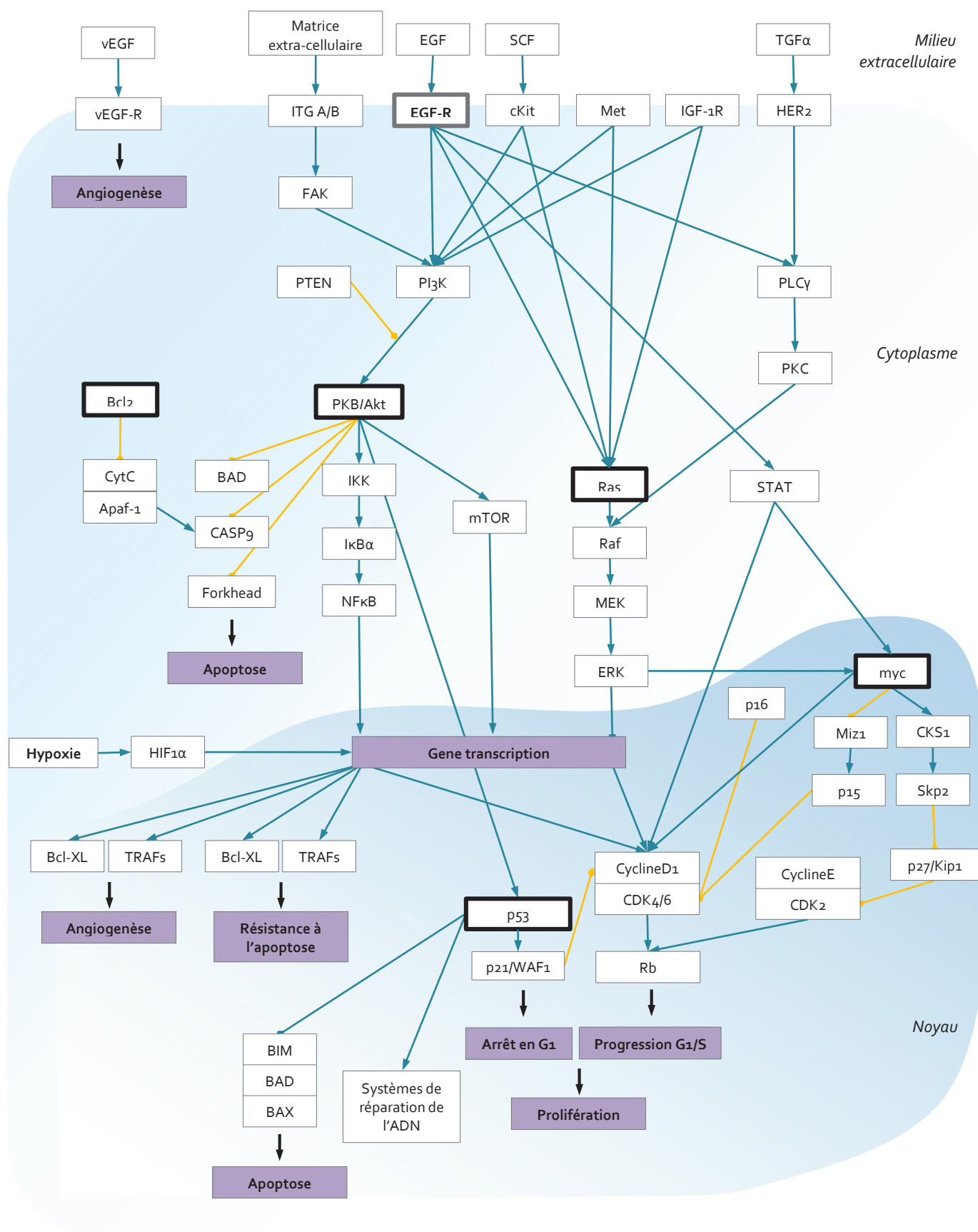


Figure 23 : Voies de signalisation couramment altérées dans les cellules cancéreuses pulmonaires

Figure réalisée à partir de (Featherstone et al., 2012 ; West et al., 2012), (Allison, 2010) et de (Sanders and Albitar, 2010).

—→ Activation —→ Répression

2.3.2.1. Altérations moléculaires promouvant la prolifération

Les cellules normales requièrent des signaux mitogènes pour passer d'un état quiescent à un état de prolifération actif. Ces signaux (facteurs de croissance diffusibles, composants de la matrice extracellulaire, molécules d'adhésion cellulaire) sont transduits via des récepteurs transmembranaires et leurs voies de signalisation en aval (**Figure 23**). Les cellules cancéreuses présentent une dépendance réduite aux signaux de croissance, qu'elles peuvent acquérir via deux mécanismes principaux : sécrétion de leurs propres facteurs de croissance et mise en place d'une boucle de régulation autocrine stimulant la prolifération, ou hypersensibilité aux signaux de croissance via des altérations des récepteurs transmembranaires (surexpression, activation constitutive ou ligand-indépendante) ou des altérations des voies intracellulaires de prolifération en aval ([Fedi et al., 1997](#)). Ainsi, on compte parmi les oncogènes pulmonaires de nombreux facteurs de croissance (comme les protéines de la famille FGF) et des récepteurs transmembranaires (dont les membres de la famille HER). Il peut ensuite s'agir de protéines impliquées dans la transduction des signaux prolifératifs : des petites protéines G, d'autres protéines membranaires liant le GTP (les plus connus étant les proto-oncogènes de la famille *RAS* fréquemment mutés dans les adénocarcinomes bronchiques) ou des protéines tyrosine-kinases (comme la protéine Src surexprimée dans les CBPC). Certains oncogènes agissent enfin comme facteurs de transcription promouvant la prolifération, comme ceux de la famille *MYC* ([Sato et al., 2007](#)).

Par exemple, les protéines HER (HER-2 et HER-1/EGFR) sont des récepteurs transmembranaires des signaux de prolifération (TGF α et EGF respectivement) à activité tyrosine kinase et à rôle oncogène. La protéine HER-2 est surexprimée dans 10% à 30% des CBNPC et plus particulièrement dans les adénocarcinomes peu différenciés. L'augmentation du nombre de récepteurs HER-2 à la surface d'une cellule cancéreuse permet l'activation de la croissance cellulaire par la voie PLC γ /PKC (**Figure 23**), conduisant à une survie diminuée chez les patients concernés ([Tan et al., 2003](#)). De la même manière, le récepteur membranaire au facteur de croissance épidermique EGF (EGFR ou HER-1) est détecté dans 40% à 70% des CBNPC ([Sato et al., 2007](#)), plus particulièrement dans les carcinomes bronchiques épidermoïdes. En présence de son ligand et suite à l'auto-phosphorylation de son domaine tyrosine kinase, ce récepteur enclenche la voie de prolifération cellulaire p13K/AKT/mTOR ([Gustafson et al., 2010](#)) (**Figure 23**).

Dans un tissu normal, des signaux antiprolifératifs (externes et internes à la cellule) assurent le maintien de la quiescence cellulaire et de l'homéostasie tissulaire, souvent via le blocage du cycle cellulaire au point de contrôle de la phase G₁. Une des caractéristiques des cellules cancéreuses est leur insensibilité à ces signaux antiprolifératifs pouvant conduire à l'acquisition d'une prolifération incontrôlée, notamment du fait de la perte de la fonction de gènes suppresseurs de tumeur et d'altérations dans les voies de signalisation intracellulaires menant aux facteurs de contrôle du cycle (complexes cyclines - kinases associées aux cyclines (CDK)). Les suppresseurs du tumeur impliqués peuvent ainsi être des régulateurs transcriptionnels qui répriment l'expression de gènes essentiels à la progression du cycle cellulaire ou d'inhibiteurs des protéines du cycle (P16-INK4A, RB) ou de facteurs impliqués dans la détection des dommages de l'ADN et assurant le blocage du cycle cellulaire tant que ces dommages ne sont pas réparés (p53) (**Figure 23**). Par exemple, la phosphoprotéine nucléaire Rb participe au contrôle du cycle cellulaire : sous sa forme non phosphorylée (active), elle participe au maintien des cellules en phase G₁ par inhibition du facteur de transcription E2F, alors que la forme phosphorylée par le complexe Cycline D1/CDK4 (inactive) autorise le passage G₁/S (**Figure 23**) ([Coqueret, 2002](#)). L'abolition de la fonction suppresseur de tumeur de Rb, courante dans les cancers pulmonaires, participe donc également à l'acquisition d'une prolifération non contrôlée ([Kaye, 2002](#)).

2.3.2.2. Altérations moléculaires inhibant la mort cellulaire

Le phénomène de mort cellulaire programmée, ou apoptose, assure la disparition d'une cellule donnée, suite à des signaux pro-apoptotiques extérieurs ou internes à la cellule. Ce processus physiologique permet notamment d'éliminer les cellules présentant trop d'aberrations ou une infection. La capacité à

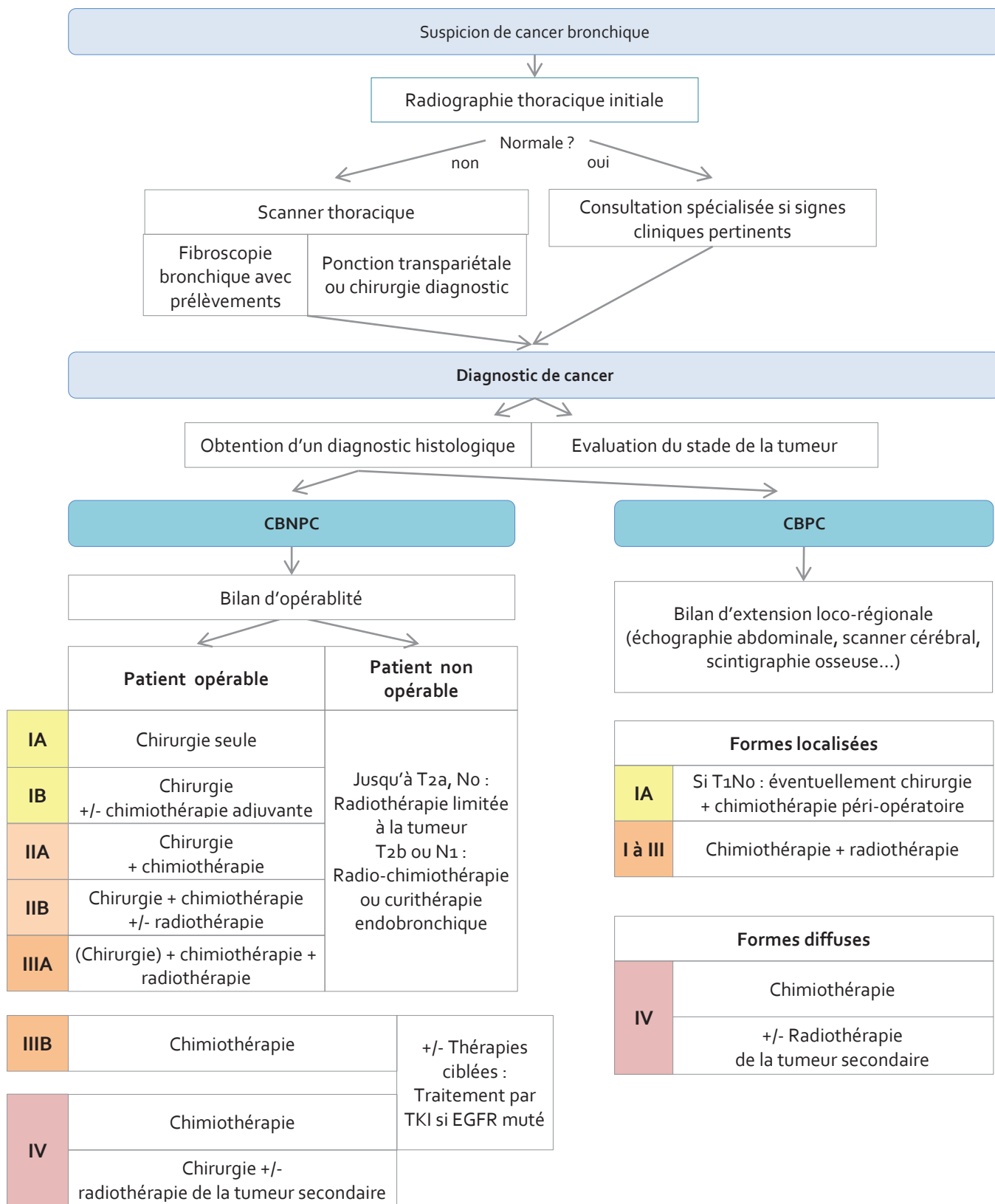


Figure 24 : Référentiel de prise en charge et de traitement des cancers bronchiques

CBPC : carcinome bronchique à petites cellules ; CBNPC : carcinome bronchique non à petites cellules.

Ce diagramme présente de manière simplifiée les standards thérapeutiques couramment utilisés dans le traitement des carcinomes bronchiques en fonction des stades d'extension tumorale.

Il faut rappeler que 60% des malades sont diagnostiqués tardivement (c'est-à-dire aux stades IIIB ou IV) et que 80% ne peuvent bénéficier d'une chirurgie.

Adapté de (INCa, 2011).

s'évader de l'apoptose est une propriété importante des cellules cancéreuses. Par ce moyen, elles acquièrent la capacité de continuer à se diviser malgré l'accumulation de défauts, sans enclencher le processus protecteur de mort cellulaire programmée. Les dérégulations dans les voies pro- et anti-apoptotiques sont courantes dans les cellules tumorales pulmonaires : l'inactivation de l'apoptose est impliquée dans la carcinogenèse bronchique, mais aussi dans les phénomènes de résistance aux traitements (Han and Roman, 2010). On a décrit la manière dont des altérations du suppresseur de tumeur p53 participent au phénomène d'échappement à l'apoptose (**voir paragraphe 2.3.1.1**). La protéine Bcl-2 est elle aussi directement impliquée dans le contrôle de la mort cellulaire programmée. Cette protéine mitochondriale est capable d'inhiber l'apoptose (Gross, 2001) en empêchant notamment l'homodimérisation des protéines pro-apoptotiques Bax et le relargage du cytochrome-c, un activateur des caspases pro-apoptotiques (**Figure 23**). Or, une surexpression de la protéine Bcl-2 est observée dans 75 à 95% des CBPC et 10 à 35% des CBNPC (Ben-Ezra et al., 1994) et la protéine agit alors comme un oncogène : la surexpression de Bcl-2, en balançant l'équilibre Bcl-2/Bax en sa faveur, aura pour conséquence une inhibition de l'apoptose et une augmentation de la survie cellulaire. (Strasser et al., 1997).

Ainsi, après un demi-siècle d'avancées rapides, la recherche sur le cancer du poumon a généré un socle important de connaissances, révélant que cette pathologie prend avant tout ses sources dans les altérations du génome. Les fondations ont été posées avec la découverte des mutations produisant des oncogènes (gain de fonction dominant) et désactivant des gènes suppresseurs de tumeurs (perte de fonction récessive). Toutefois, alors que l'étude des mutations génétiques du cancer a longtemps occupé le devant de la scène, on a assisté à l'explosion récente des connaissances sur le rôle prépondérant joué par des modifications épigénétiques et l'altération des patrons d'expression génique dans le déclenchement et la progression des cancers pulmonaires. Ces dérégulations dans l'expression des gènes feront l'objet de la troisième partie de cette introduction.

2.4. Prise en charge et traitements des tumeurs pulmonaires

2.4.1. Prise en charge initiale : détection et diagnostic

Les signes et les symptômes du cancer du poumon sont souvent discrets et tardifs - c'est entre 40 et 75 ans que la maladie est le plus souvent découverte. Les signes cliniques devant attirer l'attention, bien que communs à des affections pulmonaires bénignes, incluent respiration sifflante, souffle court ou dyspnée, toux tenace et quinteuse non soulagée par les traitements habituels (éventuellement associée à des crachements de sang ou hémoptysie), douleurs thoraciques et infections broncho-pulmonaires à répétition dans le même territoire. L'implication d'organes voisins peut provoquer dysphonie (voix rauque), dysphagie, amaigrissement, syndrome d'obstruction de la veine cave supérieure, troubles cardiaques (Minna and Schiller, 2008). Lorsque le développement tumoral est déjà bien avancé, on peut noter des symptômes associés à des métastases distantes. Enfin, les cancers peuvent causer divers syndromes paranéoplasiques, tels que les syndromes de Pierre Marie ou de Schwartz-Bartter, essentiellement dans les cancers à petites cellules. Le cancer du poumon peut être détecté par une radiographie thoracique (mais celle-ci s'avère normale dans 5% des cas de cancer pulmonaire) (**Figure 24**). Le diagnostic est confirmé par une biopsie, réalisée par endoscopie bronchique ou par ponction transpariétale guidée par scanner (dans le cas de certaines tumeurs périphériques inaccessibles). L'analyse anatomo-pathologique de l'échantillon tumoral permettra le diagnostic histologique (CBNPC ou CBPC). Une analyse par scanner thoracique et abdominal doit permettre de rechercher l'atteinte d'autres organes (en particulier le foie et les glandes surrénales, fréquemment cibles des métastases pulmonaires) et d'évaluer l'extension ganglionnaire. Néanmoins, il faut noter que les méthodes de diagnostic dépendent du type de cancer (à petites cellules ou non), de la taille et de la localisation de la tumeur primaire, de la présence éventuelle de métastases et de l'état général du patient (Rivera et al., 2003). De ce diagnostic dépendra le choix de la stratégie thérapeutique à adopter.

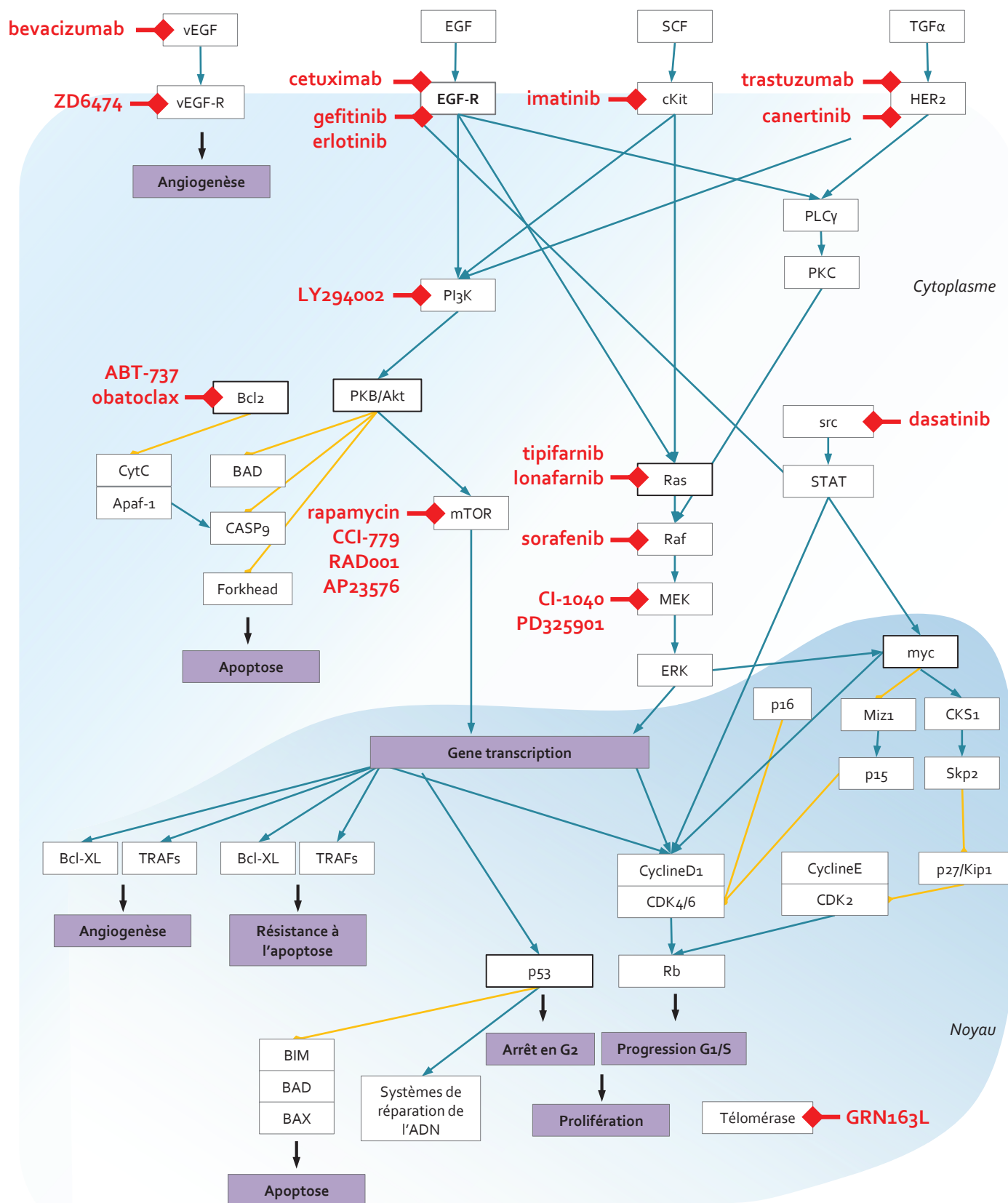


Figure 25 : Thérapies émergentes ciblant des composants des voies de signalisation impliquées dans les cancers pulmonaires.

Adapté de (Sato et al., 2007).

—→ Activation

—→ Répression

—◆ Inhibition médicamenteuse

La plupart des tumeurs pulmonaires sont détectées à un stade tardif, limitant la possibilité d'utiliser la chirurgie et la radiothérapie comme traitements. En conséquence, il apparaît primordial de développer de nouvelles méthodes de détection précoce. Des approches par CT (computed tomography)-scan hélicoïdal donnent l'espoir d'améliorer la détection précoce mais restent débattues (Black et al., 2007). On cherche aussi à développer des techniques de détection précoce non-invasives basées sur l'analyse des fluides biologiques (sang, sérum, urine, expectorations ou air expiré), à la recherche de biomarqueurs de tumeur qui peuvent être de l'ADN circulant libéré par les cellules cancéreuses ou des anticorps associés aux tumeurs dans le sérum, des aneusomies chromosomiques dans les crachats, des adduits sur l'ADN dûs à l'exposition aux carcinogènes du tabac dans l'urine, ou encore l'augmentation de certains composés organiques volatils dans l'air expiré reflétant le métabolisme des cellules cancéreuses (D'Urso et al., 2013 ; Mascaux et al., 2010).

2.4.2. Les traitements conventionnels et leurs limites

Il existe trois leviers thérapeutiques conventionnels pour le traitement des cancers bronchiques : la chirurgie, la radiothérapie et la chimiothérapie. Ces traitements peuvent être réalisés de manière isolée, séquentielle ou combinée. Le choix de la stratégie à adopter dépend de nombreux critères (Figure 24).

2.4.2.1. Chirurgie

La première approche qui peut être envisagée est l'exérèse chirurgicale de la tumeur. Cependant, cette stratégie thérapeutique ne s'applique quasiment jamais aux cas de CBPC, tumeurs très agressives et souvent déjà disséminées au moment du diagnostic. De nombreux facteurs sont à prendre en compte avant chirurgie d'un CBNPC : l'extension tumorale telle qu'elle est définie par la classification TNM, la notion de résécabilité et d'opérabilité ou encore l'état général du patient (pas de risque cardiovasculaire identifié, réserve respiratoire pré-opératoire suffisante). Une chimiothérapie pré-opératoire peut permettre d'obtenir une réduction du volume tumoral, ce qui facilite ensuite le geste chirurgical (Betticher, 2005). La chirurgie peut consister en une résection périphérique (une partie d'un lobe), une segmentectomie (une division anatomique d'un lobe), une lobectomie (un lobe), une bilobectomie (deux lobes) ou la pneumonectomie (poumon entier) (Collins et al., 2007). La première pneumonectomie réussie comme traitement d'un cancer du poumon a été réalisée dès 1933 (Horn and Johnson, 2008). L'opération chirurgicale en elle-même a un taux de décès d'environ 4,4 %. Néanmoins, on estime que plus de 80% des cancers broncho-pulmonaires ne peuvent pas être opérés à cause d'un diagnostic trop tardif. Même pour les CBNPC où la chirurgie est le traitement de référence, elle n'est possible que dans 25% des cas, lorsque l'atteinte est limitée à un seul poumon, et seulement jusqu'au stade IIIA (Figure 24).

2.4.2.2. Radiothérapie

La radiothérapie palliative (de faibles doses de radiation pour contrôler la douleur) a été utilisée depuis les années 1940 ; elle est toujours utilisée aujourd'hui lors de la présence de métastases douloureuses comme les métastases osseuses (Ishiyama et al., 2004). La radiothérapie radicale (à haut intensité) a été utilisée dès les années 1950 ; et en 1997 apparaît la radiothérapie continue hyperfractionnée accélérée, où une forte dose de radiothérapie est appliquée sur un temps bref. Souvent associée à la chimiothérapie, la radiothérapie utilise des radiations ionisantes (rayons X et électrons produits par les accélérateurs linéaires) qui ciblent l'ADN, conduisant à l'arrêt de prolifération et/ou la mort des cellules tumorales. La curiethérapie est une variante qui vise à appliquer la radiothérapie localement et à haute dose, directement dans l'arbre bronchique quand le cancer bloque une grosse voie aérienne, ou après une chirurgie sur les lèvres d'une résection pour abaisser le taux de récurrence. Il faut noter que la radiothérapie ne pourra être utilisée que sur les formes de carcinome bronchique limitées au thorax (30% des cas). Le traitement local par radiothérapie sera obsolète sur les formes disséminées où seule la chimiothérapie sera utilisée. Une radiothérapie peut néanmoins être réalisée sur la tumeur secondaire (métastase) si son extension et sa localisation le permettent (Figure 24).

2.4.2.3. Chimiothérapie

La chimiothérapie anticancéreuse est un mode de traitement reposant sur l'utilisation de drogues ciblant plus ou moins spécifiquement les cellules cancéreuses. La majorité des substances chimiothérapeutiques classiques affectent les cellules à fort indice prolifératif, en interférant avec les processus mitotiques. Les CBPC disséminés sont traités uniquement par chimiothérapie (**Figure 24**). Une chimiothérapie primaire est aussi utilisée pour les CBNPC non opérables ou déjà métastasés. Les premières combinaisons de chimiothérapies efficaces pour le traitement des carcinomes bronchiques ont été mises au point dans les années 1970 ([Cohen et al., 1977](#)). La chimiothérapie actuelle pour les CBNPC est généralement basée sur une administration de cisplatine et/ou de drogues de 3^{ème} génération (vinorelbine, paclitaxel, docétaxel et gemcitabine) ([Clegg et al., 2002](#)). Les CBPC quant à eux présentent une bonne chimiosensibilité initiale. Le cisplatine et l'étoposide sont fréquemment utilisés ([Murray and Turrisi, 2006](#)). Les programmes associant plusieurs drogues permettent d'obtenir des taux de réponse complète aux alentours de 50 à 60% dans les formes localisées, 15 à 45% dans les formes disséminées. Néanmoins, 90 % de ces cancers nécessiteront un traitement de deuxième ligne, du fait d'une non-réponse à la première ligne, d'une progression après réponse partielle ou d'une récurrence après réponse complète ([Fennell, 2003](#)). Des combinaisons de carboplatine, gemcitabine, paclitaxel, lomustine, vinorelbine, topotécan et irinotécan seront alors utilisées ([Azim and Ganti, 2007](#) ; [MacCallum and Gillenwater, 2006](#)). Les CBPC, considérés comme fortement chimio- et radio-sensibles, sont en réalité rarement chimiocurables. La médiane de survie des patients porteurs d'une forme disséminée ne dépasse pas 10 mois et peu de patients sont en vie deux ans après le diagnostic.

2.4.3. Traitements émergents : les thérapies ciblées

Malgré les progrès observés au cours de ces vingt dernières années dans le traitement du cancer bronchique, en particulier avec la chimiothérapie à base de sels de platine, les bénéfices en termes de survie (notamment dans les formes avancées de CBPC) restent modestes et semblent avoir atteint un plateau. Néanmoins, les avancées dans la caractérisation moléculaire des tumeurs pulmonaires, la compréhension des voies cellulaires impliquées dans le processus de transformation néoplasique (**voir paragraphe 2.3.3**) et le développement de la pharmacogénétique, ont permis l'émergence de nouvelles drogues ([Sun et al., 2007b](#)). Ces traitements, ciblant spécifiquement des protéines et des voies de signalisation identifiées dans les cancers bronchiques (**Figure 25**), assurent une meilleure sélectivité thérapeutique et doivent permettre d'atténuer les effets toxiques des thérapeutiques moins sélectives comme la chimiothérapie conventionnelle.

De nombreuses études ont été menées à ce jour, visant à cibler spécifiquement l'une des étapes de la transduction du signal de prolifération : les récepteurs HER, l'étape de farnésylation de RAS, la voie PI3K/PTEN/AKT/mTOR, les protéines kinase C (PKC) ou encore les cyclines kinase-dépendantes (CDK). Par exemple, deux approches ont été développées pour bloquer l'activation des récepteurs HER : d'une part la production d'anticorps monoclonaux humanisés dirigés contre leur domaine extracellulaire, dont le cetuximab (pour EGF-R) et le trastuzumab (pour HER-2) ; d'autre part le développement d'inhibiteurs des domaines tyrosine kinase (TKI) tels que le gefitinib et l'erlotinib (**Figure 26**). Ces deux derniers composants ont démontré leur aptitude à accroître le taux de survie des patients souffrant de CBNPC ([Feld et al., 2006](#)). Par ailleurs, certains patients CBNPC présentent des mutations somatiques dans le domaine TK de l'EGFR (10 à 55% des cas d'adénocarcinomes) ([Sequist et al., 2007](#)). Ces mutations (mutations ponctuelles ou petites délétions/insertions conservant le cadre de lecture) dans les exons 19 ou 21 principalement, causent une activation constitutive de l'EGF-R ([Greulich et al., 2005](#)). Les cellules tumorales développent alors une forte dépendance à ce récepteur constitutivement actif (un phénomène appelé « addiction à l'oncogène ») et sont particulièrement sensibles à l'inhibition de son activité, alors que les cellules normales sont peu affectées. Ainsi, le taux de réponse à un traitement TKI est supérieur à 60% chez les patients possédant un EGF-R muté, contre 10% pour les cas négatifs pour cette mutation ([Sequist et al., 2007](#)). En conséquence, les traitements par TKI sont particulièrement

indiqués chez les patients présentant un EGF-R muté et la recherche de ces mutations constitue un élément important dans le choix de la stratégie thérapeutique à adopter (**Figure 25**).

La caractérisation moléculaire des tumeurs et la prise en compte de la multiplicité des sous-types moléculaires fait ainsi passer d'une thérapie ciblée à une thérapie personnalisée ([Ogino et al., 2012](#)). Les approches de profilage moléculaire (**voir paragraphe 2.3.3**), dont le but ultime serait d'identifier tous les changements moléculaires présents dans la tumeur de n'importe quel patient, pourraient être utilisées pour la détection moléculaire précoce, la prédiction du comportement biologique de la tumeur et du pronostic clinique, et le développement et la sélection de thérapeutiques adaptées. Mais malgré l'amélioration récente des indications opératoires, notamment grâce au progrès des techniques d'imagerie et la diffusion de traitements innovants, aucune amélioration franche de la survie n'a été observée entre 1989-1991 et 2001-2004 en France pour les carcinomes bronchiques ([Grosclaude et al., 2013](#)).

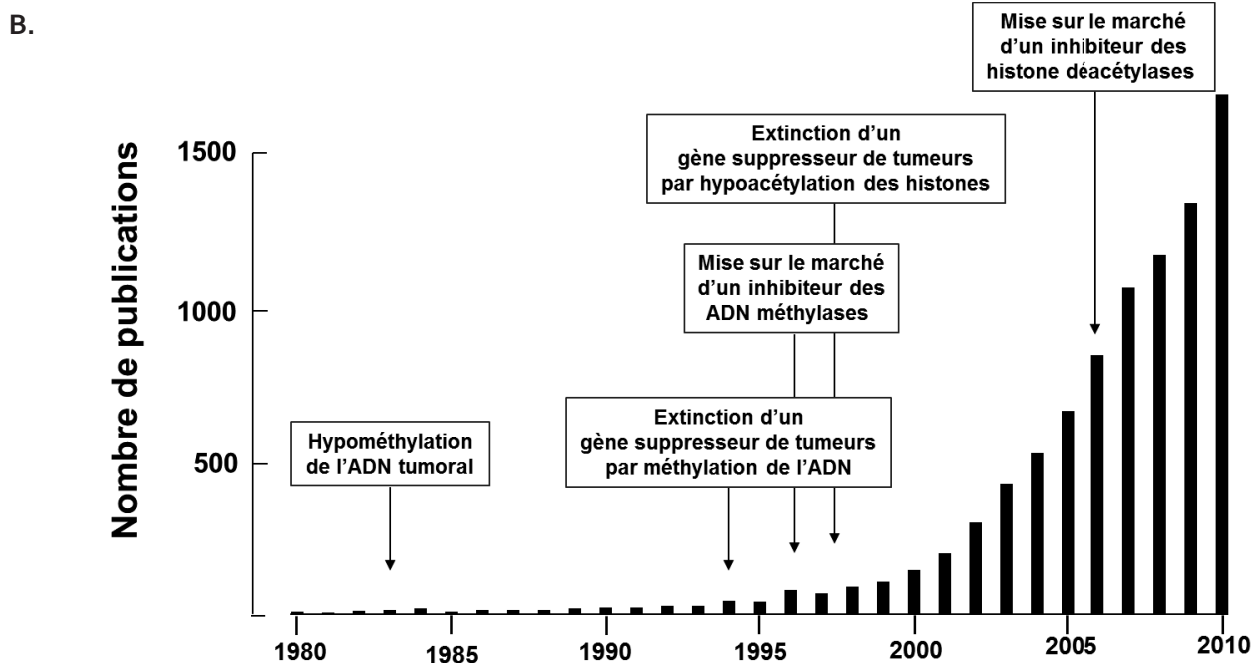
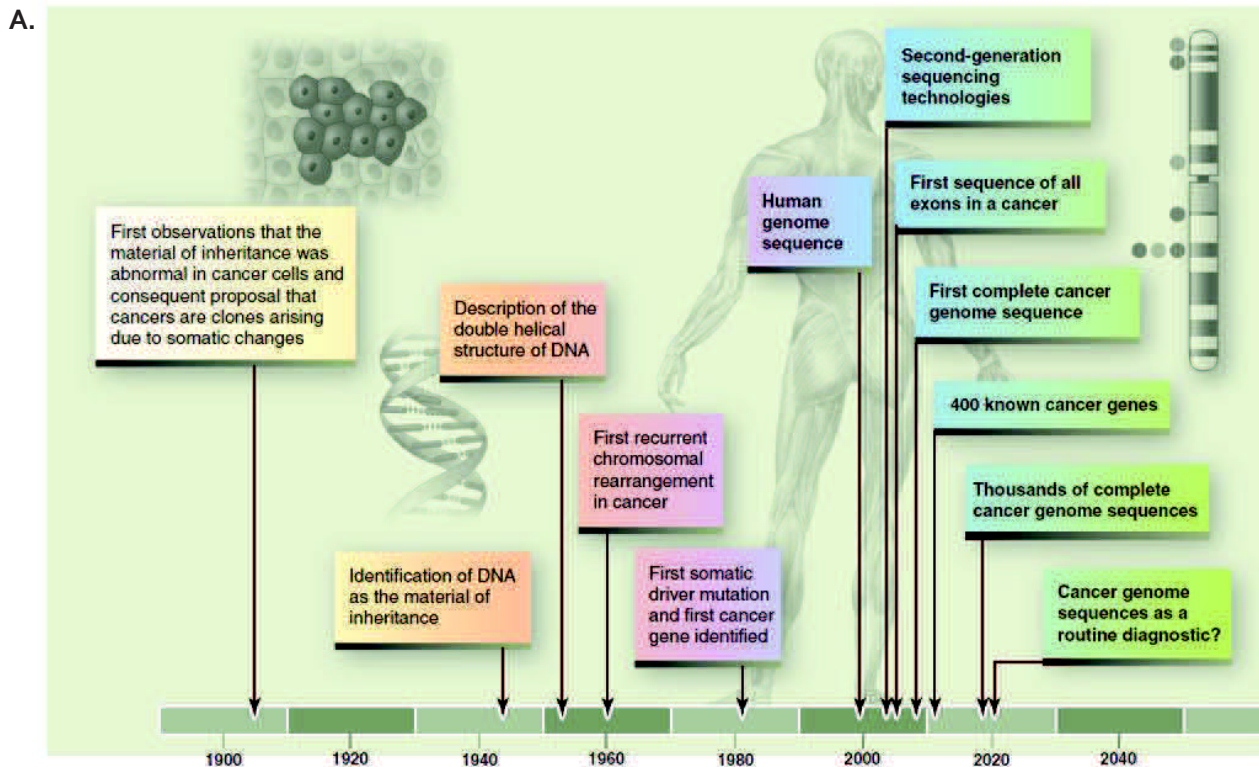


Figure 26 : La compréhension des causes génétiques et épigénétiques dans les cancers

A. Frise chronologique présentant les principaux évènements et découvertes dans le champ « Génome et cancer ».

Tiré de (Stratton, 2011).

B. Évolution du nombre de publications scientifiques et découvertes majeures dans le domaine « Épigénétique et cancer ».

Source : Pubmed.

3. ALTERATIONS DES EXPRESSIONS GENIQUES DANS LES CELLULES CANCEREUSES PULMONAIRES

3.1. Altérations génétiques et épigénétiques globales dans les cellules cancéreuses

Nous avons présenté dans la seconde partie de cette introduction quelques exemples choisis de mutations ponctuelles liées à la carcinogenèse pulmonaire - les mutations de *TP53* liées au tabac par exemple (voir paragraphe 2.3.1.2). Mais les cellules cancéreuses n'exhibent pas que quelques gènes-clés mutés : c'est l'ensemble du génome qui accumule des altérations, à la fois génétiques et épigénétiques, au cours de la progression tumorale. Alors que la littérature s'est cristallisée au cours des années passées sur les altérations génétiques subies par quelques candidats seulement (oncogènes et gènes suppresseur de tumeur qui ont été intensivement étudiés), on a pris conscience plus récemment de l'ampleur et du caractère global des perturbations affectant le génome des cellules cancéreuses (Figure 26A). La partie qui suit va tenter de donner un aperçu de ces altérations génétiques et épigénétiques et d'en comprendre les causes.

3.1.1. Vers l'accumulation progressive d'altérations génétiques

3.1.1.1. Mutations ponctuelles : théorie de la mutagenèse somatique

Pendant les années, voire dizaines d'années que dure le tabagisme de l'individu, les cellules de l'épithélium bronchique sont la cible des agents mutagènes de la fumée de cigarette, et leur ADN est notamment soumis à la formation continue d'adduits. S'ils ne sont pas réparés, ces adduits sont à l'origine de mutations ponctuelles somatiques qui s'accumulent dans le génome au fil du temps (voir paragraphe 2.1.2.1) (Hang, 2010 ; Pfeifer et al., 2002). Certaines de ces mutations, affectant notamment les oncogènes et gènes suppresseurs de tumeur, peuvent conférer un avantage à la cellule cancéreuse qui les porte, comme une capacité de division accrue ou une plus grande résistance aux signaux de mort cellulaire (voir paragraphe 2.3.2) : ces mutations seront donc sélectionnées au cours du développement tumoral. Cette situation est résumée par la théorie de la mutagenèse somatique, énoncée pour la première fois par Thilly en 1983 (Thilly, 1983) puis modifiée en 1997 (Holmquist and Gao, 1997). Celle-ci théorise le lien existant entre les substances génotoxiques et la transformation maligne en établissant que la probabilité qu'une mutation se retrouve dans une tumeur se formule de la manière suivante :

$$P_i(\text{mutation dans la tumeur}) = P_i(\text{endommagée}) \times P_i(\text{non réparée}) \times P_i(\text{répliquée}) \times P_i(\text{sélection}).$$

Ainsi, la probabilité P_i qu'une base à la position i dans le génome soit mutée dans une tumeur dépend de la probabilité de plusieurs événements qui interviennent successivement. Il faut que cette position nucléotidique soit effectivement endommagée par l'agent mutagène, que ce dommage ne soit pas réparé par les systèmes cellulaires de réparation des dommages à l'ADN et que la base endommagée soit lue et mal répliquée par l'ADN polymérase. Enfin, pour que la mutation s'étende à la majorité des cellules composant une tumeur, il faut que cette mutation ait été sélectionnée pour l'avantage qu'elle confère (Holmquist and Gao, 1997). Lorsque qu'une cellule est sélectionnée pour une mutation qui lui confère un avantage (mutation « driver ») et qu'elle se divise activement, elle transmet à ses cellules-filles cette mutation qui se propage alors au sein de la tumeur, mais aussi toutes celles qu'elle a accumulées en d'autres points du génome (mutations passagères).

Au fil du développement de la tumeur, les cellules cancéreuses accumulent donc des mutations, jusqu'à exhiber un génome complètement altéré (Hubaux et al., 2012). Les techniques d'analyse à « haut débit » du génome, notamment le séquençage massif, permettent aujourd'hui de recenser et de cartographier les mutations ponctuelles. Ces travaux ne cessent de démontrer l'ampleur des mutations somatiques dans les cancers du poumon. Une première étude a reporté les cas de mutations ponctuelles dans la séquence codante

de plus de 500 protéines kinase dans 200 types de tumeurs humaines. Les cancers du poumon sont décrits comme l'un des types de cancer comptant le plus grand nombre de ces mutations somatiques (4.21 mutations en moyenne par mégabase d'ADN) (Greenman et al., 2007). Une seconde étude, du projet « Tumor Sequencing Project (TSP) » s'est focalisée sur les adénocarcinomes pulmonaires en séquençant le génome de 188 d'entre eux (Ding et al., 2008). Sur un total de 247 mégabases analysées, on a dénombré 1 013 mutations somatiques non synonymes, dont 915 mutations ponctuelles et 12 mutations dinucléotides. Au total, cette étude a identifié 26 gènes mutés de manière significative, dont plusieurs oncogènes et gènes suppresseurs de tumeur déjà décrits (*TP53*, *KRAS*, *CDKN2A*, *STK11/LKB1*...). Enfin, deux travaux publiés en 2010 ont produit l'analyse la plus détaillée des altérations génétiques dans le cancer du poumon (Lee et al., 2010 ; Pleasance et al., 2010). Par séquençage du génome complet d'une tumeur pulmonaire, les auteurs ont comptabilisé respectivement 22 910 et plus de 50 000 mutations somatiques, absentes du génome d'un tissu non cancéreux séquençé pour comparaison. La plupart d'entre elles étaient silencieuses, mais respectivement 134 et 530 identifiées dans des séquences codantes (dans les gènes *TP53*, *KRAS* et *RB* notamment) pourraient jouer un rôle majeur dans le processus de carcinogenèse.

3.1.1.2. Mutabilité accrue et phénotype d'instabilité génétique

L'ampleur des altérations génétiques dans les cellules cancéreuses pulmonaires peut s'avérer surprenante, comparé aux taux attendus de mutations induites par les agents mutagènes du tabac (Loeb, 1991). La présence dans une cellule cancéreuse d'un nombre élevé de mutations sur le génome peut en partie s'expliquer par le caractère clonal de la croissance tumorale : une cellule qui acquiert un avantage lié à une mutation se divise activement et transmet cette mutation à ses cellules -filles ; si une nouvelle mutation s'opère chez l'une des cellules-filles, les deux mutations seront co-sélectionnées et se diffuseront ensemble dans la population des cellules tumorales (voir paragraphe 2.3.1). Ainsi, le spectre de mutations de *p53* trouvées dans les tumeurs pulmonaires est considéré non seulement comme une signature de l'exposition aux mutagènes de la fumée de cigarette (hydrocarbures aromatiques polycycliques notamment), mais également comme le reflet de la sélection tumorigénique (Rodin and Rodin, 2000 ; Rodin and Rodin, 2005).

Certains ont suggéré que les cellules cancéreuses devaient, en plus de cela, avoir acquis une mutabilité accrue (Loeb, 1991). L'acquisition de ce « phénotype mutateur » peut être expliquée par l'altération des composants gardiens et réparateurs du génome (Lengauer et al., 1998). En réponse à un dommage au génome, la protéine *p53* agit comme un senseur des dommages, puis assure normalement l'arrêt du cycle cellulaire pour permettre les réparations de l'ADN, ou dirige la cellule vers l'apoptose si les dommages sont trop importants. Quand la fonction de cette protéine est perdue au cours de la progression tumorale, comme c'est le cas dans plus de la moitié des cancers pulmonaires (Sato et al., 2007), cela autorise l'accumulation progressive de mutations. D'autre part, un événement mutagène qui viendrait altérer un des composants des systèmes de réparation de l'ADN aurait également pour conséquence une accumulation accrue de mutations dans le génome, étant donné que les défauts générés par les agents mutagènes du tabac ne pourraient plus être corrigés. On a ainsi pu discuter de l'implication des mutations de l'ADN glycosylase OGG1. Si cette enzyme est mutée, la perte de sa fonction ne permet plus d'assurer la réparation des 8-oxoguanines générées par la fumée de cigarette (Abedin et al., 2013), ce qui favoriserait l'apparition de nouvelles mutations. L'expression de la protéine MGMT (O6-méthylguanine-DNA méthyltransférase), une enzyme de réparation de l'ADN assurant la suppression des dommages de type méthylation des guanines, est couramment réprimée dans les adénomes pulmonaires (Pulling et al., 2003), un événement pouvant là aussi participer à l'accumulation d'altérations génétiques. Enfin, une susceptibilité accrue au cancer pulmonaire a été décrite chez les patients présentant un polymorphisme dans certains gènes des systèmes de réparation (Hsia et al., 2012 ; Yin et al., 2007 ; Yu et al., 2008). Ainsi, les défauts acquis au cours de la progression tumorale dans les systèmes de détection et de réparation de l'ADN permettent l'apparition d'une instabilité génétique dans les cellules cancéreuses, instabilité qui favorise la survenue de nouveaux événements mutagènes et l'émergence de cellules cancéreuses mutantes pouvant présenter de nouveaux avantages sélectifs (Loeb, 2011). Cette instabilité génétique est considérée par

Hanahan et Weinberg comme la caractéristique majeure (« enabling characteristic ») permettant l'émergence des six caractéristiques fondamentales des cellules cancéreuses (voir paragraphe 2.3.2 et Figure 22) (Hanahan and Weinberg, 2011).

Il faut signaler que la déficience dans les systèmes de réparation de l'ADN, qui constitue un des moteurs de la carcinogenèse pulmonaire, peut être retournée contre la tumeur. Les cellules cancéreuses présentant des défauts dans les voies de réparation seraient beaucoup plus sensibles à une chimiothérapie par un agent induisant des dommages à l'ADN (Postel-Vinay et al., 2012). D'un point de vue clinique, une déficience dans les systèmes de détection ou de réparation des dommages à l'ADN serait associée à une tumeur à évolution rapide et un pronostic initial plus sombre pour le patient, mais constituerait un marqueur prédictif de bonne réponse aux traitements. Le composant ERCC1, une endonucléase impliquée dans la réparation des adduits de platine par le système de réparation par excision de nucléotides (nucleotide excision repair NER), est actuellement l'un des candidats les plus prometteurs : une déficience de ERCC1 dans les CBNPC semble associée à une sensibilité accrue au traitement par le cisplatine (Olaussen et al., 2006).

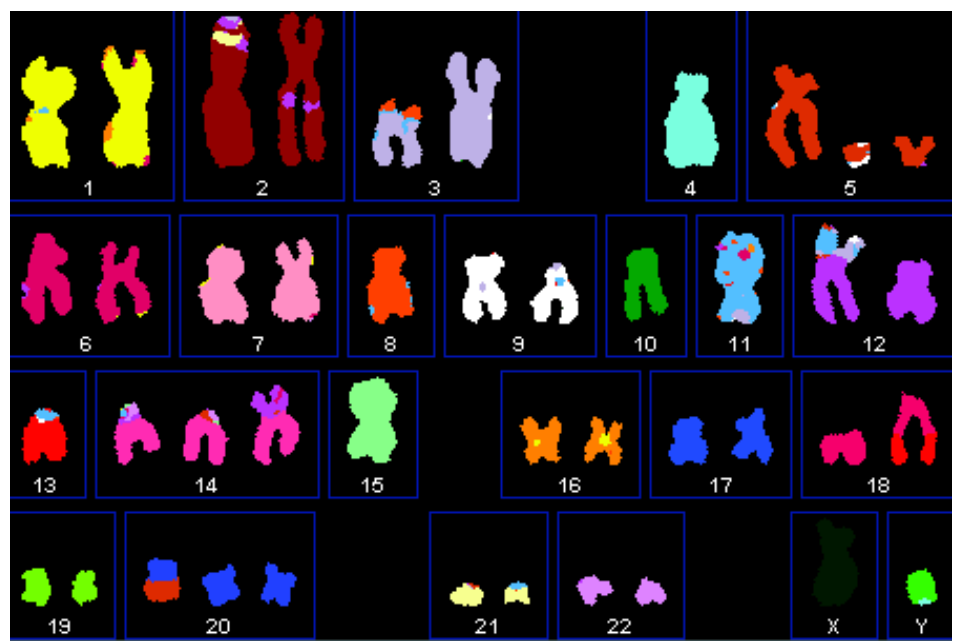
3.1.1.3. Instabilité chromosomique et altérations de grande ampleur

Les altérations du génome des cellules cancéreuses pulmonaires acquises au cours de l'évolution de la tumeur ne se limitent pas aux mutations ponctuelles, mais peuvent être de grande ampleur. La présence d'aberrations chromosomiques a été documentée dès la fin des années 1980 pour les CBPC (Naylor et al., 1987). A l'analyse anatomo-pathologique, les cellules cancéreuses biopsées présentent de fréquentes anomalies cytogénétiques souvent complexes (Testa et al., 1994) : anomalies quantitatives (hypoploïdie, hyperploïdie, polyploïdie, aneuploïdie, trisomies ou monosomies variables) et anomalies qualitatives (délétions, translocations, inversions), parfois associées à des anomalies sévères de la morphologie des chromosomes (chromosomes en anneaux ou dicentriques, isochromosomes). Ces anomalies sont visibles sur un simple caryotype, mais sont aujourd'hui couramment analysées par les techniques de FISH (Hybridation in situ en fluorescence - *Fluorescent in situ Hybridization*) ou de SKY (caryotype spectral - *Spectral Karyotype*) grâce à des sondes couplées à des fluorochromes (Figure 27A). Les analyses des défauts structuraux des chromosomes ont notamment permis d'identifier des événements courants dans les carcinomes bronchiques, comme les pertes de 3p et 9q. Ainsi, le petit bras du chromosome 3 présente des délétions dans plus de 90% des CBPC et 60 à 80% des CBNPC (Whang-Peng et al., 1982). On a identifié plus récemment un événement de translocation qui peut être détecté dans 5% des CBNPC, surtout chez les non-fumeurs ; il s'agit d'une inversion sur le chromosome 2, conduisant à la fusion EML4-ALK (Soda et al., 2007) (voir paragraphe 3.3.4).

Les analyses de type CGH array (comparative genomic hybridization) permettent quant à elles d'étudier plus précisément les réarrangements chromosomiques causant le gain ou la perte de matériel chromosomique. L'étude de 2010 précédemment évoquée a conduit cette analyse CGH sur un génome de tumeur pulmonaire et a identifié 43 variations structurales de grande ampleur causant des déséquilibres alléliques, des pertes d'homozygoties ou des variations du nombre de copies (copy number variants CNV), incluant des variations déjà connues (perte de *TP53*, gain de *KRAS* ou perte d'hétérozygotie de *RB*) (Lee et al., 2010). Une étude plus anciennes, menée sur 371 adénocarcinomes pulmonaires avait identifié 57 événements récurrents conduisant à des variations du nombre de copies, dont seulement une poignée avaient auparavant été associées à un gène précis (Weir et al., 2007). Ainsi, ce type d'analyses, en plus de mettre en lumière l'étendue des altérations chromosomiques, permet d'identifier de nouveaux candidats, oncogènes ou suppresseurs de tumeurs, qui seraient impliqués dans la tumorigenèse pulmonaire.

L'ensemble de ces anomalies chromosomiques peuvent être détectées dans les cellules cancéreuses, mais sont également plus fréquentes dans les cellules bronchiques saines des individus gros fumeurs (Nelson et al., 1998 ; Varella-Garcia et al., 2007 ; Wistuba et al., 1997). Certaines d'entre elles, comme les pertes de 3p et de 9q, restent détectables dans les cellules de l'épithélium bronchique plusieurs années après l'arrêt de l'intoxication tabagique (Mao et al., 1997 #307).

A.



B.

Zones de pertes alléliques	Gènes impliqués	Fonctions
3p14.2	FHIT	Apoptose
3p21.3	Sémaphorine, β -caténine	Adhésion, arrêt en G1
3p24.2	RAR β	Différenciation
9p21	P16 ^{INK4}	Arrêt en G1
5q21	APC-MCC	Adhésion
13q14	Rb	Arrêt en G1
17p13	p53	Arrêt en G1, apoptose

Figure 27 : Perte de matériel génétique et gènes suppresseurs de tumeur associés dans les cellules cancéreuses pulmonaires

A. Réarrangements chromosomiques dans la lignée cancéreuse pulmonaire NCI-H21711 hypodiploïde (perte de 3p, 4, gain de 20, 5 dicentrique, translocations...) observés par la technique de caryotype spectral (SKY).

Tiré de (Grigorova et al., 2005).

B. Localisation des principaux gènes dont les délétions alléliques semblent impliquées dans la carcinogenèse bronchique. *D'après (Brambilla et al., 1996).*

Les mécanismes à l'origine de ce phénomène d'instabilité chromosomique (chromosome instability - CIN) sont peu à peu élucidés. Les anomalies du nombre de chromosomes sont principalement dues à des ségrégations anormales du matériel génétique lors des divisions cellulaires, du fait d'altérations du point de contrôle du fuseau mitotique, de la formation de centrosomes anormaux ou de défauts de cytokinèse (Masuda and Takahashi, 2002). Les altérations chromosomiques structurales seraient plus particulièrement causées par des défauts de réparation des cassures double-brin de l'ADN (*DNA double-strand breaks* - DSB) du fait d'altérations dans les systèmes de réparations de ces DSB (van Gent et al., 2001) ou dans les points de contrôle sensés détecter les dommages à l'ADN. Le rôle de l'érosion des télomères dans l'acquisition du phénotype d'instabilité chromosomique a également été suggéré. En effet, le raccourcissement intensif auquel les télomères sont soumis du fait de la prolifération des cellules cancéreuses pourrait favoriser la formation de chromosomes dicentriques ou en anneaux (Artandi and DePinho, 2000 ; Harley and Sherwood, 1997).

Quelle que soit l'échelle considérée, le génome des cellules cancéreuses pulmonaires présente donc une forte instabilité génétique, depuis l'accumulation de milliers de mutations ponctuelles jusqu'à des réarrangements chromosomiques massifs et fréquents. Etant donné l'ampleur de ces altérations génétiques, ce phénotype d'instabilité se double d'altérations dans les expressions géniques. Ainsi, un gène qui s'exprime en temps normal dans le poumon sain, mais qui présente une délétion dans sa séquence codante ou au niveau de son promoteur, pourra voir son expression réduite ou abolie dans la tumeur (voir paragraphe 3.2.1.1). A l'inverse, un gène présentant un gain de nombre de copies via une amplification génique pourra être surexprimé (voir paragraphe 3.2.2). Enfin, les translocations, en créant des gènes fusion ou en déplaçant un gène normalement silencieux derrière un promoteur fort, pourront être à l'origine d'expressions aberrantes de gènes dont les expressions sont normalement réprimées (voir paragraphe 3.3.4).

3.1.2. Les altérations épigénétiques

Nous avons vu qu'au sein de la chromatine, l'information génétique portée par la séquence de l'ADN se double d'une information épigénétique portée par les modifications dynamiques et réversibles des cytosines de l'ADN et des queues d'histones (voir paragraphe 1.1.3). Le paysage épigénétique est finement régulé dans une cellule saine, afin d'assurer d'une part l'expression d'un répertoire restreint de gènes essentiels à la cellule, et d'autre part la répression à grande échelle de gènes qui doivent rester silencieux, assurant de fait le maintien de l'identité cellulaire. Dans les cellules cancéreuses, ce paysage épigénétique est totalement bouleversé (Esteller, 2008 ; Jones and Baylin, 2007). Cette compréhension de l'ampleur des bouleversements épigénétiques dans le cancer est relativement récente, comme en témoigne les occurrences du terme « épigénétique et cancer » dans les publications scientifiques au cours des dernières années (Figure 26B). Alors que les mécanismes génétiques occupaient jusqu'alors le devant de la scène, ce n'est qu'à partir du milieu des années 1990 que l'on a pris conscience de l'importance et de l'intérêt des mécanismes épigénétiques dans le développement tumoral, avec notamment la découverte de l'extinction du premier gène suppresseur de tumeur, le gène *VHL*, par méthylation de l'ADN (Herman et al., 1994). Tout comme le paragraphe précédent a décrit de manière générale les altérations génétiques, nous allons tenter de donner ici un aperçu global des altérations épigénétiques fréquemment observées dans les cellules cancéreuses pulmonaires, avant de décrire plus en avant les conséquences sur les expressions géniques.

3.1.2.1. Défauts de méthylation de l'ADN

La méthylation de l'ADN d'une cellule cancéreuse peut être étudiée par un panel de techniques, dont certaines ont été développées récemment : technique du « shotgun bisulfite sequencing » (MethylC-seq), immunoprécipitation de l'ADN méthylé grâce à un anticorps anti-méthylcytosine (MeDIP-seq), capture de l'ADN méthylé par purification par affinité (MethylCap-seq), séquençage des régions d'ADN associées aux domaines MBD (MBD-seq) ou encore utilisation d'enzymes de restriction sensibles à la méthylation de leur site de coupure (MRE-seq) (Figure 28A) (Bock et al., 2010 ; Harris et al., 2010). Ces techniques ont permis de

démontrer que les patrons de méthylation de l'ADN des cellules cancéreuses exhibent un double défaut, à savoir des hyperméthylations localisées associées à une hypométhylation globale.

3.1.2.1.1. Hyperméthylations localisées

La première observation dans les cancers pulmonaires d'une hyperméthylation localisée date de la fin des années 1980 : une hyperméthylation du gène de la calcitonine, un marqueur du CBPC, fut mise en évidence et corrélée à son extinction transcriptionnelle (Baylin et al., 1986). De fait, l'hyperméthylation en des loci particuliers est une caractéristique majeure des altérations épigénétiques dans les cancers pulmonaires (Divine et al., 2005 ; Zochbauer-Muller et al., 2002) généralisable à l'ensemble des cancers (Esteller, 2007). Ces hyperméthylations touchent en particulier les îlots CpG au sein de promoteurs, notamment de gènes suppresseurs de tumeur, et sont associées à une répression transcriptionnelle (Herman and Baylin, 2003). Il y a des preuves solides que ces hyperméthylations gène-spécifiques et les répressions de gènes suppresseurs de tumeur qui en découlent constituent une étape fondamentale dans le développement des tumeurs (Esteller, 2007). Les hyper-méthylations localisées sont en effet détectées dès les lésions précurseurs bénignes (Belinsky et al., 1998) et passent pour être des événements précoces de la cancérogenèse pulmonaire (Suzuki and Yoshino, 2010). De nombreux gènes « ciblés » par cette méthylation aberrante des promoteurs ont été décrits et sont associés à des fonctions variées, telles que l'arrêt du cycle cellulaire (p16, p53), la réparation de l'ADN (MGMT) ou la limitation de la prolifération (RASSF1A, APC, FHIT, PTEN) (Heller et al., 2010). Certains d'entre eux seront décrits plus en détails lorsque l'on abordera les répressions aberrantes de gènes (voir paragraphe 3.2.1.2).

Certains suggèrent de plus que l'hyperméthylation de l'ADN pourrait participer à l'instabilité génétique des cellules cancéreuses en les prédisposant aux événements mutagènes (notamment par déamination spontanée des cytosines méthylées) (Van Den Broeck et al., 2010). Les cytosines méthylées pourraient ainsi être à l'origine de 30% des mutations de type transition trouvées dans les cancers (Tornaletti and Pfeifer, 1995). La méthylation des CpG favoriserait également la fixation du benzo(a)pyrène de la fumée de cigarette, créant ainsi des points chauds (hotspots) mutationnels, un phénomène qui a été particulièrement étudié dans le cas de *TP53* (Yoon et al., 2001).

On a rapidement envisagé que ces hyperméthylations locales puissent être le fait d'altérations des effecteurs épigénétiques responsables de la méthylation des CpG. Les protéines DNMT1, 3a et 3b sont exprimées à des hauts niveaux dans les carcinomes pulmonaires, en particulier chez les fumeurs (Lin et al., 2005). De plus, ces surexpressions corrélaient positivement avec l'hyperméthylation des promoteurs des gènes suppresseurs de tumeur et sont associées à une diminution de la survie (Kim et al., 2006a ; Lin et al., 2007b ; Xing et al., 2008). Récemment, on a même fait le lien entre consommation de tabac et surexpression de DNMT1 : *in vivo*, le carcinogène du tabac 4-(methylnitrosamino)-1-(3-pyridyl)-1-butanone (NNK) limite la dégradation de DNMT1 et promeut son accumulation dans le noyau, favorisant de fait la méthylation de l'ADN (Lin et al., 2010 ; Damiani et al., 2008 ; Liu et al., 2010).

3.1.2.1.2. Hypométhylation globale

La diminution globale du taux de méthylcytosine de l'ADN a été la première aberration épigénétique identifiée dans les cellules cancéreuses au début des années 1980 (Feinberg et al., 1988 ; Feinberg and Tycko, 2004) et l'on estime qu'une cellule saine compte deux fois plus de résidus cytosine méthylés qu'une cellule cancéreuse (Dunn, 2003). Lorsque l'on considère le génome dans son ensemble, les cellules cancéreuses pulmonaires présentent une hypométhylation globale dont les causes restent relativement méconnues (Esteller and Herman, 2002). Etant donné que près de la moitié de notre génome est constitué de séquences hautement répétées normalement méthylées, la baisse générale du niveau de méthylation pourrait être imputée à la déméthylation de ces régions (Pfeifer and Rauch, 2009). La déméthylation des régions répétées pourrait de plus augmenter la possibilité de réarrangements chromosomiques, participant de fait à l'acquisition du phénotype

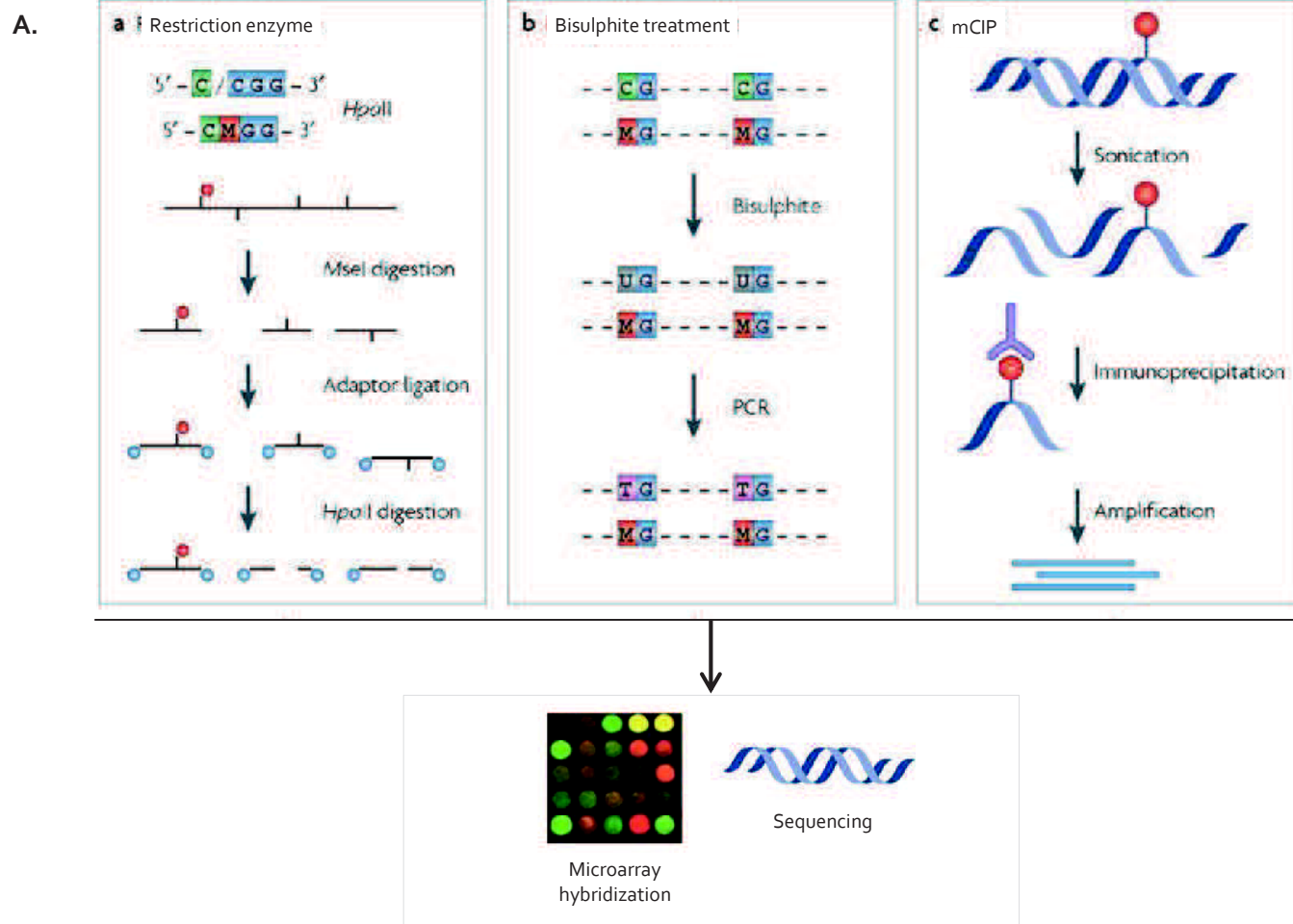
d'instabilité génétique (Eden et al., 2003). A l'inverse des hyperméthylations localisées, la déméthylation globale est considérée comme un événement plutôt tardif de la carcinogenèse pulmonaire (Selamat et al., 2011). Certains ont également suggéré l'existence d'hypométhylations régionales au niveau de certains promoteurs ou d'éléments transposables (Chilukamarri et al., 2007 ; Lin et al., 2007a ; Szyf et al., 2004) qui pourraient conduire à l'activation de certains proto-oncogènes, à la perte de l'empreinte ou à la réactivation d'éléments transposables.

3.1.2.2. Défauts de modifications des histones

Les modifications post-traductionnelles des histones dans les cellules cancéreuses peuvent être étudiées à l'échelle du génome par immunoprécipitation de la chromatine avec des anticorps spécifiques des différents types de modifications (ChIP), couplée à une hybridation sur puce (ChIP-chip) ou à un séquençage haut-débit (ChIP-seq) (Barski et al., 2007 ; Schones and Zhao, 2008).

De la même manière que pour la méthylation de l'ADN, on a pu mettre en évidence des altérations de modifications des histones dans les cancers en général (Chi et al., 2010) et dans les cancers pulmonaires en particulier. Ainsi, les cellules cancéreuses bronchiques exhibent une diminution des niveaux d'acétylation de H3K18, H4K12, H4K16 et de méthylation de H3K4 et H4K20 associée à une augmentation de l'acétylation de H4K5 et H4K8 et de la triméthylation de H3K27 (Li et al., 2011 ; Van Den Broeck et al., 2008). Certains de ces changements (baisse de H4K20me3) apparaissent dès le stade des lésions précancéreuses et progressent de manière graduelle au cours de la carcinogenèse, d'autres (baisse de H3K4me2, de H3K18ac et de H4K20me3) ont été corrélées à un pronostic plus sombre pour les patients (Seligson et al., 2009 ; Van Den Broeck et al., 2008). En conséquence, l'analyse des modifications d'histones pourraient aider au diagnostic des tumeurs pulmonaires (Barlesi et al., 2007). Des études *in vitro* soulignent l'implication du tabac dans ces modifications aberrantes des histones. Ainsi, des cellules épithéliales saines des voies aériennes cultivées en présence de condensats de fumée de cigarette accumulent de manière temps- et dose-dépendante des changements dans les modifications des histones caractéristiques de la progression tumorale pulmonaire (diminution de H4K16Ac et H4K20Me3, augmentation de H3K27Me3) (Liu et al., 2010).

Ces défauts de modifications des histones pourraient tout d'abord être directement liés aux hyperméthylations aberrantes. En effet, une des caractéristiques des protéines à domaine MBD se fixant sur les CpG méthylés est de pouvoir recruter des complexes ayant des activités de modifications post-traductionnelles des histones entraînant une extinction transcriptionnelle (voir paragraphe 2.1.2.2.3) (Sarraf and Stancheva, 2004). Dans les cellules cancéreuses pulmonaires, l'hyperméthylation des CpG pourraient donc entraîner l'apposition et/ou le retrait aberrants de modifications épigénétiques des histones. D'autre part, on a suggéré que ces défauts de marques d'histones seraient dus à des altérations des effecteurs épigénétiques qui mettent en place (HAT, HMT) ou qui effacent (HDAC, HDM) ces marques. Concernant les histones méthyltransférases, on a observé que la perte de H4K20me3 corrèle avec une diminution d'expression de l'HMT SUV4-20H2 (Van Den Broeck et al., 2008) et il s'avère que la dérégulation des HMT participe directement à la transformation oncogénique des cellules bronchiques (Watanabe et al., 2008). Concernant les histones acétyltransférases, la baisse de transcription du gène codant la HAT MYST Tip60 pourrait expliquer l'hypoacétylation de H4K16 (L. Leonart et al., 2006). Des mutations ou délétions du gène codant la HAT CBP pourraient également favoriser une baisse du niveau d'acétylation des histones (Kishimoto et al., 2005). Les éléments cette fois-ci liés aux « effaceurs » épigénétiques telles que les histones déacétylases ou déméthylases sont plus parcellaires, et les corrélations entre leur niveau d'expression, leur activité et la progression tumorale sont variables. Par exemple, l'expression de HDAC1 semble augmenter au cours de la progression tumorale (Sasaki et al., 2004), HDAC3 est surexprimée dans les carcinomes épidermoïdes comparé au tissu pulmonaire non-tumoral (Bartling et al., 2005) alors que les HDAC de classe II présentent une baisse d'expression corrélée à un pronostic sombre (Osada et al., 2004). Quant à HDAC2, fréquemment surexprimée dans les tumeurs solides, elle semble conférer un potentiel oncogénique aux cellules pulmonaires, en dérégulant l'expression de protéines de l'apoptose (p53, Bax, Bcl-2)



B.

Gènes avec promoteur hyperméthylé	LOH	Fonctions
P16 ^{INK4}	oui	Cycle cellulaire
RARβ	oui	
RASSF1	oui	
Rb	oui	
FHIT	oui	
DAPK		Apoptose
O ⁶ MGMT		Réparation de l'ADN
E cadhérine	oui	Migration et adhésion
SEMA3B		
APC		

Figure 28 : Hyperméthylation des promoteurs de gènes suppresseurs de tumeur dans les cellules cancéreuses pulmonaires

A. Les méthodes d'étude de la méthylation de l'ADN. Tiré de (Schones and Zhao, 2008).

a. Utilisation d'enzymes de restriction sensibles à la méthylation de leur site de coupure (Ex : *HpaII*).

b. Traitement de l'ADN au bisulfite qui transforme les cytosines non méthylées en uridines.

c. Immunoprécipitation des fragments d'ADN méthylé grâce à un anticorps spécifique.

Chacune de ces méthodes peut ensuite être combinée à une hybridation sur puce ou à du séquençage.

B. Principaux gènes suppresseurs de tumeur exhibant une hyperméthylation du promoteur lors de la carcinogenèse pulmonaire.

LOH : perte d'hétérozygotie. D'après (Kohno and Yokota, 1999) et (Van Den Broeck et al., 2010).

et du cycle cellulaire (p21/WAF-1, cyclines E2 et D1, CDK2) (Jung et al., 2012). Enfin certaines HDM, dont JMJD2C, semblent être surexprimées dans les cancers bronchiques (Italiano et al., 2006). L'ensemble de ces résultats confirme ainsi que des modifications épigénétiques aberrantes du génome pourraient être impliquées dans l'apparition du cancer, dans sa progression aussi bien que dans son agressivité.

3.2. Dérégulations des expressions géniques

3.2.1. Inactivations aberrantes de gènes suppresseurs de tumeur

Dans la seconde partie de cette introduction, nous avons illustré comment la fonction suppresseur de tumeur de la protéine p53 pouvait être perdue du fait de mutations ponctuelles liées au tabac dans la séquence génique de *TP53* (voir paragraphe 2.3.1.2). Néanmoins, une substitution dans la séquence protéique n'est pas l'unique processus conduisant à l'inactivation d'un facteur suppresseur de tumeur. Très souvent, les altérations ont lieu au niveau de l'expression du gène, ces dérégulations d'expressions géniques étant la conséquence directe des altérations génétiques et épigénétiques du génome des cellules bronchiques. La perte de l'expression d'un gène suppresseur de tumeur s'explique par deux mécanismes principaux : la perte physique de la séquence génique ou de séquences régulatrices du fait d'une délétion (Figure 27B) ou la répression aberrante de l'expression du gène du fait d'une hyperméthylation du promoteur (Figure 28B).

3.2.1.1. Pertes physiques des loci suppresseurs de tumeurs

Les pertes physiques de gènes suppresseurs de tumeur ont été suggérées par l'observation de délétions récurrentes dans les tumeurs pulmonaires, dans des régions connues pour héberger certains gènes suppresseurs de tumeur. Ainsi, dans la majorité des cas de CBPC où le gène *TP53* n'est pas fonctionnel, on observe une délétion allélique au niveau du locus 17p13 (Lee et al., 2010 ; Weston et al., 1989). De la même manière, le petit bras du chromosome 3 présente de fréquentes pertes alléliques (dans plus de 90% des CBPC et 60 à 80% des CBNPC) (Kok et al., 1997) et l'établissement de cartographies détaillées des locus 3p ont peu à peu permis d'identifier de nouveaux candidats comme gènes suppresseurs de tumeurs (Zabarovsky et al., 2002). Le gène *FHIT* (fragile histidine triad) par exemple (Sozzi et al., 1996), localisé en 3p14.2, est fréquemment délété de manière homozygote et son expression est perdue dans 50% des cancers pulmonaires (Sozzi et al., 1998 ; Sozzi et al., 1997). On a pu montrer que *FHIT* est un facteur qui limite la prolifération cellulaire (Siprashvili et al., 1997), qui induit l'apoptose (Ji et al., 1999 ; Sard et al., 1999) en lien avec la voie des caspases (Roz et al., 2002) et dont la réexpression réduit la tumorigénicité *in vivo* (Ji et al., 1999 ; Roz et al., 2002), ce qui fait de *FHIT* un bon candidat suppresseur de tumeur dans les cancers pulmonaires. Dans leurs ensemble, les délétions de gènes suppresseurs de tumeur ont été extensivement étudiées dans les cancers du poumon (An et al., 2002 ; Kohno et al., 2010), et l'on a pu démontrer de fréquentes pertes d'hétérozygoties corrélées à des pertes d'expression des protéines correspondantes pour des gènes tels que *TP53*, *FHIT*, *IFNA*, *CDKN2A/p16* ou *RB* notamment (Figure 27B) (Gouyer et al., 1994).

3.2.1.2. Répressions aberrantes des expressions de suppresseurs de tumeur

En plus des événements génétiques, des événements épigénétiques vont également être impliqués dans les répressions aberrantes de l'expression de certains gènes suppresseurs de tumeur. Le gène *CDKN2A* (Cyclin-dependent kinase inhibitor 2A) est localisé au niveau de la région 9p21 et la perte d'un allèle est observée dans environ 50% des cancers du poumon. On observe aussi des mutations de l'un des allèles dans environ 15% des tumeurs pulmonaires (Okamoto et al., 1994). Mais dans un tiers des carcinomes bronchiques, c'est un événement d'hyperméthylation du promoteur de *CDKN2A* qui va réprimer l'expression de ce gène. Ainsi, la combinaison de perte d'hétérozygotie voire de délétion homozygote, de mutation et d'hyperméthylation de *CDKN2A* conduit à l'inactivation complète du gène, qui est bien corrélée à la perte de la protéine (Brambilla et al., 1999 ; Gazzeri et al., 1998). Le gène *CDKN2A* code la protéine p16^{INK4}, un inhibiteur de kinase cycline-

dépendante (CDK) qui joue un rôle essentiel dans le contrôle du cycle via le blocage de la liaison du couple CyclineD1-CDK4/6. La conséquence de la perte de p16 est une expression accrue de la cycline D1 qui conduit à une hyperphosphorylation persistante de Rb et à l'emballement du cycle cellulaire (**Figure 23**) (Gautschi et al., 2007). L'hyperméthylation de *CDKN2A* est associée à un pronostic sombre (Lou-Qian et al., 2013). Elle est considérée comme un évènement précoce de la carcinogenèse bronchique et sa fréquence augmente au cours de la progression tumorale (Belinsky et al., 2005 ; Belinsky et al., 1998 ; Nuovo et al., 1999). De fait, l'hyperméthylation aberrante de *CDKN2A* s'avère être un biomarqueur précoce de cancer du poumon. Elle peut être détectée dans les expectorations des patients à carcinomes épidermoïdes, et ceci jusqu'à trois ans avant le diagnostic clinique de cancer pulmonaire (Palmisano et al., 2000).

Cette hyperméthylation des séquences régulatrices a été démontrée pour de nombreux autres gènes suppresseurs de tumeur (**Figure 28B**). L'application de nouvelles technologies haut-débit à l'étude des CBNPC (Gomperts et al., 2011) a permis d'identifier l'hyperméthylation de gènes tels que *RARβ* (Retinoic Acid Receptor Beta) dans 40–43% des cas, *O⁶MGMT* dans 16 à 27% ou encore *RASSF1A* (*RAS association domain family 1A*) dans 30 à 40% (Zochbauer-Muller et al., 2002). De manière intéressante, on a montré que les hyperméthylations de *RARβ*, *p16*, *FHIT* et *RASSF1A* sont corrélées au statut de fumeur du patient et que le degré de méthylation augmente avec l'intensité du tabagisme (Talikka et al., 2012). Comme cela est le cas pour *CDKN2A*, on estime que l'hyperméthylation de ces promoteurs pourrait constituer un marqueur précoce pour aider à la détection des tumeurs pulmonaires (Gomperts et al., 2011 ; Masciaux et al., 2010). Ainsi, l'hyperméthylation des promoteurs *CDKN2A*, *O⁶MGMT*, *APC* ou *RASSF1A* peut être détectée dans le sang (Esteller et al., 1999) ou les expectorations (Belinsky et al., 2002) de patients atteints de tumeurs pulmonaires. On a également montré que l'on peut utiliser l'ensemble de ces hyperméthylations pour préciser le pronostic d'une tumeur. Individuellement, de nombreuses hyperméthylations sont des marqueurs de mauvais pronostic, comme celles de *RASSF1A*, *PTEN*, *H-cadherin* ou *APC* (Adenomatous Polyposis coli) (Wen et al., 2011) et la combinaison de l'hyperméthylation de quatre gènes (*CDKN2A*, *CDH13*, *RASSF1A* et *APC*) chez les patients CBNPC de stade I est associée à une récurrence précoce (Brock et al., 2008).

3.2.1.3. Vers des thérapies épigénétiques ?

Au contraire des mutations somatiques qui modifient de manière définitive la séquence génique, le caractère épigénétique des hyperméthylations des gènes suppresseurs de tumeur leur confère une réversibilité. Cette réversibilité pourrait se manifester spontanément *in vivo* après l'arrêt de l'intoxication tabagique, étant donné que les anciens fumeurs ont un indice de méthylation de l'ADN plus faible que les fumeurs actuels (Yanagawa et al., 2011b). On peut aussi obtenir la déméthylation des CpG via des traitements déméthylants, en particulier par l'analogue de nucléosides 5-aza-2'-deoxycytidine (5AzaCdR), inhibiteur de DNMT1. De nombreuses études sur cellules cancéreuses pulmonaires en culture ont montré qu'un traitement à la 5AzaCdR permet de supprimer l'hyperméthylation aberrante des promoteurs et de restaurer l'expression de certains gènes suppresseurs de tumeur (Goffin and Eisenhauer, 2002 ; Miki et al., 2001). Ces réexpressions restaurent la fonction suppressive des gènes suppresseur de tumeur, comme c'est le cas de la réexpression de *RASS1A* dans des cellules de CBPC en culture qui conduit à une croissance cellulaire ralentie et qui inhibe la formation de tumeur dans des souris nude (Burbée et al., 2001). Cliniquement, il a été montré qu'un traitement par la 5AzaCdR augmente la survie de patients atteints de CBNPC (Juergens et al., 2011 ; Momparler et al., 2000), mais ces résultats restent discutés. En particulier, on a pu craindre que les analogues de nucléotides augmentent l'instabilité génomique et le risque de développer d'autres cancers (Gaudet et al., 2003 ; Jackson-Grusby and Jaenisch, 1996 ; Jones and Taylor, 1980), d'où le développement d'inhibiteurs de DNMT non-analogues de nucléosides (Castellano et al., 2011 ; Yu and Wang, 2008). Enfin, il faut souligner que, même si les hyperméthylations semblent jouer un rôle important dans la carcinogenèse pulmonaire et constituent des cibles thérapeutiques prometteuses, il reste que les "règles" qui gouvernent la manière dont ces profils de méthylation sont établis au niveau de loci spécifiques durant la carcinogenèse sont encore inconnues.

Une autre catégorie de drogues épigénétiques se révèle pouvoir cibler les répressions aberrantes de gènes. Il s'agit des inhibiteurs d'histone déacétylases (HDACi), capables de maintenir un haut niveau d'acétylation des queues d'histones. Il semble en effet que la surexpression de certaines HDAC et les déacétylations inappropriées puissent contribuer à la répression transcriptionnelle de gènes suppresseurs de tumeur et à la progression tumorale (Carew et al., 2008). Les HDACi peuvent être de classes structurales variées, dont des acides gras à courte chaîne (comme le butyrate) ou des acides hydroxamiques (comme la trichostatine A (TSA)) capables de se lier à l'atome de zinc du site catalytique des HDAC (Marks et al., 2000). On a observé que certains HDACi autorisent la réexpression de gènes suppresseurs de tumeur précédemment réprimés (Brehm et al., 1998 ; Wu et al., 2008) et ont une activité antitumorale dans les CBNPC (Mukhopadhyay et al., 2006) notamment par induction de l'apoptose (Choi, 2005). Enfin, certaines études *in vitro* ont souligné que les inhibiteurs de DNMT et de HDAC peuvent agir de manière synergique pour induire la ré-expression de gènes impliqués dans l'apoptose, la différenciation ou l'arrêt du cycle. Par exemple, le phénylbutyrate et la 5AzaCdR utilisés en combinaison ont des effets synergiques sur la réduction de la croissance tumorale pulmonaire dans des modèles de souris (Belinsky et al., 2003).

3.2.2. Surexpressions d'oncogènes

Les facteurs oncogènes peuvent être activés par des mutations ponctuelles dans la séquence génique du proto-oncogène, comme c'est le cas pour le gène *KRAS* (voir paragraphe 2.3.1.1). Mais certaines protéines à activité oncogénique ne sont pas mutées et constitutivement activées, mais seulement surexprimées. Ces surexpressions peuvent découler directement des altérations génétiques et épigénétiques observées dans le génome des cellules cancéreuses bronchiques. Le proto-oncogène *ERBB2/HER-1* codant le récepteur à l'EGF-R (voir paragraphe 2.3.2.1) présente une amplification génique dans 10 à 20% des carcinomes bronchiques, conduisant à l'augmentation du nombre de copies de la séquence nucléotidique du proto-oncogène sur le génome disponibles à la machinerie de transcription. Parfois, il y a même polysomie du chromosome 7 porteur de *HER-1* (Hirsch et al., 2003). Tout comme les mutations de l'EGF-R (voir paragraphe 2.4.3), l'augmentation du nombre de copies de *HER-1* (détectée par FISH) et la forte expression de la protéine EGF-R (en immunohistochimie) dans les CBNPC sont des marqueurs d'un pronostic sombre, mais peuvent être utilisées comme prédicteurs de bonne réponse aux TKI (Zhu et al., 2008). De la même manière, le proto-oncogène *ERBB2* codant le récepteur HER2 (Human Epidermal Growth Factor Receptor 2) (voir paragraphe 2.3.2.1) est surexprimé du fait d'une amplification génique dans 10% à 30% des CBNPC, plus particulièrement dans les adénocarcinomes peu différenciés. Cette amplification génique corrèle avec le niveau d'expression de la protéine HER-2, mais également avec une survie diminuée chez les patients concernés (Tan et al., 2003). Les altérations du nombre de copies dans les CBNPC semblent liées au tabagisme : les gros fumeurs (plus de 60 paquets-années) exhibent davantage de gains de copies que les non- ou petits fumeurs (Huang et al., 2011).

Dans certains cancers, des événements de translocation peuvent aboutir à la relocalisation d'un proto-oncogène dans un nouveau site du génome conduisant à une augmentation d'expression. Dans certains cas de lymphomes, des translocations entre les chromosomes 8 et 14 placent les gènes *BCL2* ou *myc* à la suite du promoteur fort des chaînes lourdes des immunoglobulines, d'où une importante surexpression (Rabbitts, 1994 ; Strasser et al., 1997). Des translocations de *myc*, détectées par FISH dans des lignées primaires de cancer du poumon (Lu et al., 1996) pourraient ainsi jouer un rôle dans la carcinogenèse bronchique.

Alors que la surexpression de certains oncogènes apparaît clairement comme l'un des événements fondamentaux de la carcinogenèse bronchique, on a mis en lumière l'existence d'expressions ectopiques dans les cellules cancéreuses. Des gènes, dont l'expression est normalement restreinte à un tissu ou à un type cellulaire donné et verrouillée hors de leur tissu d'expression, peuvent être exprimés de manière ectopique dans les cellules cancéreuses. C'est sur ces expressions hors-contexte que va se focaliser la fin de cette introduction.

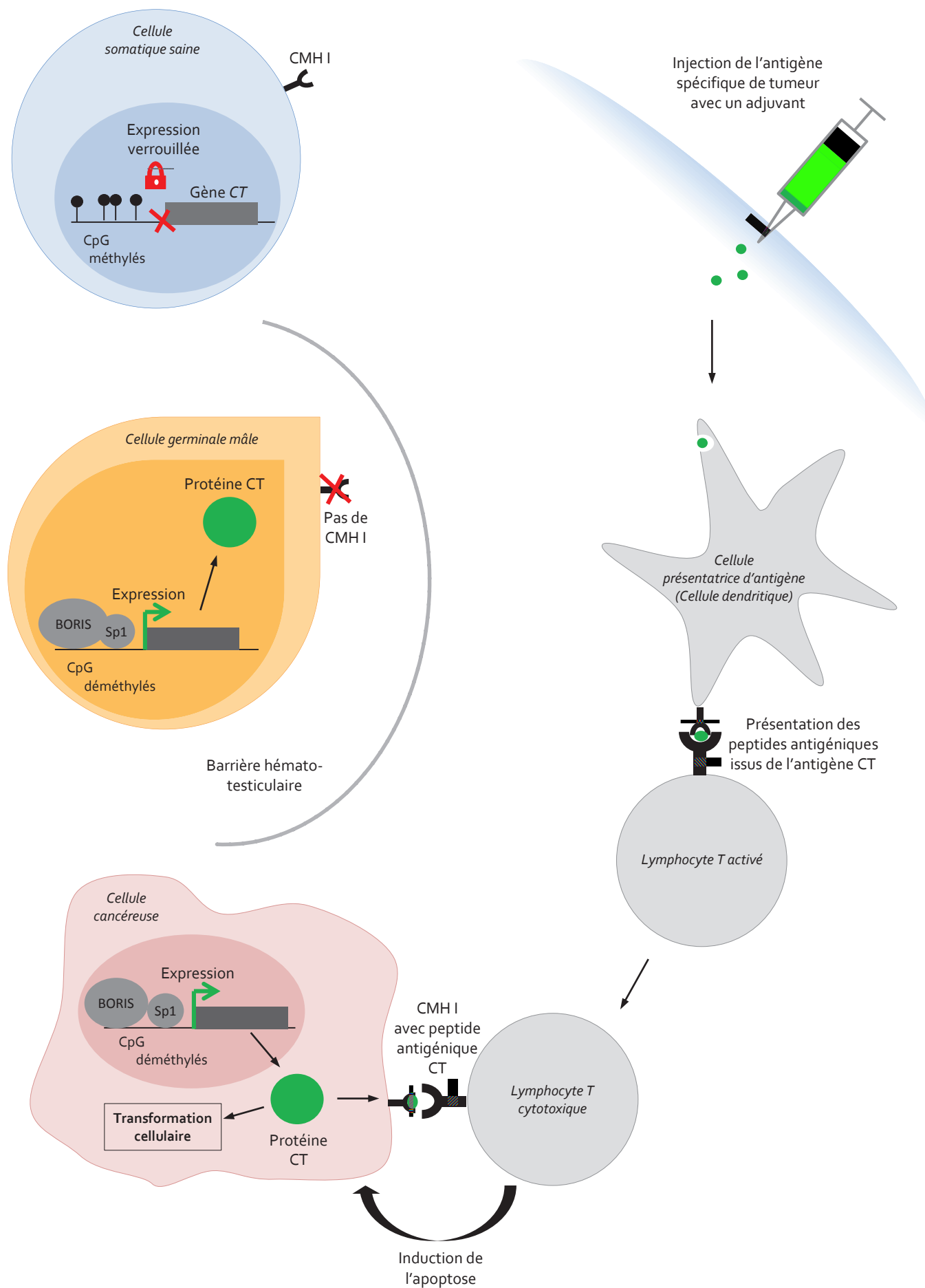


Figure 29 : Principe théorique de la vaccination anticancéreuse par un antigène spécifique de tumeur
 Inspiré de (Hiret et al., 2010).

3.3. Expressions ectopiques de gènes tissu-spécifiques

La différenciation cellulaire s'accompagne de l'établissement d'un patron d'expressions géniques, largement basé sur le contrôle du paysage épigénétique (voir paragraphe 1.3). Dans les cellules différenciées, les mécanismes épigénétiques permettent non seulement l'activation et le maintien de l'expression de gènes spécifiques du type cellulaire ou du tissu, mais contrôlent aussi la répression globale des gènes qui doivent rester silencieux, et notamment les gènes spécifiques d'autres types cellulaires (Figure 11A). Il y a quelques années, on a pourtant observé la ré-expression ectopique, dans des cellules cancéreuses somatiques, de gènes spécifiques de la lignée germinale mâle, appelés en conséquence « gènes cancer/testis » ou « gènes CT ». Ce sont les analyses de ces gènes CT dans divers types de cancers qui ont mis en lumière l'ampleur de ces expressions hors-contexte, leurs conséquences sur le phénotype cancéreux et l'intérêt de les étudier en tant que biomarqueurs ou cibles thérapeutiques.

3.3.1. Historique : la découverte des gènes CT comme cibles d'immunothérapies

La stimulation d'une réponse immunitaire dépend de la présence d'antigènes identifiés comme un corps étranger par le système immunitaire hôte. La mise au point d'un éventuel vaccin anti-cancer repose donc en théorie sur la découverte d'un antigène spécifique des cellules tumorales, pouvant être dérivé de toutes classes de protéines (enzymes, récepteurs, facteurs de transcription) mais devant être présent uniquement dans les cellules tumorales (Figure 29). Il peut donc s'agir de produits de gènes mutés, de produits provenant d'autres cadres ouverts de lecture (Open Reading Frame ORF), de produits issus d'épissage protéique spécifique aux cellules cancéreuses – ou plus intuitivement d'un produit issu d'un gène exprimé uniquement dans les cellules cancéreuses et en aucun autre endroit connu du système immunitaire. Les antigènes CT correspondent à des protéines exprimées exclusivement dans les cellules de la lignée germinale mâle (spermatocytes, spermatides, spermatozoïdes), cellules qui sont protégées par la barrière hémato-testiculaire (blood-testis barrier - BTB) et inaccessibles au système immunitaire (Figure 29) (Fijak et al., 2011). Ces protéines, exprimées ectopiquement dans des cellules cancéreuses somatiques, peuvent en théorie être détectées par les cellules immunitaires et leur antigenicité pourrait être utilisée pour développer un vaccin qui stimulerait cette réponse immunitaire anti-cellules cancéreuses pour combattre la croissance tumorale (Scanlan et al., 2002b). C'est ce qui a guidé, au cours des dernières années, la recherche sur de nombreux aspects des gènes CT, et particulièrement leur antigenicité.

Le premier antigène CT a été découvert au début des années 1990 chez une patiente porteuse d'un mélanome dont l'évolution était anormalement favorable : il s'agit du facteur *MAGE-A1* (Melanoma Antigen Gene A1), normalement absent des mélanocytes mais exprimé par les cellules de mélanome et reconnu par les lymphocytes T cytotoxiques (van der Bruggen et al., 1991). Suite au clonage de *MAGE-A1*, on s'est rendu compte que son expression était détectable au sein de plusieurs mélanomes, carcinomes du sein et autres cancers, mais jamais dans des tissus sains, à l'exception du testicule (Chen et al., 1994). Cette découverte initiale a rapidement été suivie de l'identification des antigènes CT des familles *BAGE* et *GAGE*, facteurs spécifiques du testicule mais exprimés dans les mélanomes ainsi que dans d'autres types histologiques de tumeurs (Boel et al., 1995 ; De Backer et al., 1999). Puis l'approche SEREX (serological identification of antigens by recombinant expression cloning) (Sahin et al., 1995) a mis à profit la réponse immunitaire humorale des patients pour identifier de nouveaux antigènes CT. Ainsi, l'analyse des anticorps dans le sérum de patients souffrant d'une tumeur de l'œsophage a permis la découverte du facteur *NY-ESO-1* (New York esophageal squamous cell carcinoma-1), dont l'expression est spécifique des gonades mais détectée dans un large panel de tumeurs, et qui s'est révélé être l'un des antigènes CT les plus immunogènes (Chen et al., 1997). Enfin, un certain nombre d'antigènes CT ont été identifiés sur la base de leur profil d'expression (Axelsen et al., 2007 ; Scanlan et al., 2002a). Aujourd'hui, on compte plus d'une quarantaine de familles représentant plus de 110 gènes, présentant tous la particularité d'avoir une expression restreinte au testicule et aux cancers. Ces facteurs CT identifiés jusqu'alors sont regroupés dans une base de données, « The CTDatabase » (<http://www.cta.lncc.br/>) développée par le Ludwig Institute for Cancer Research.

Ainsi, les gènes CT, dont l'expression est normalement restreinte aux cellules de la lignée germinale mâle ont un statut immunitaire tout à fait particulier. Dans des conditions physiologiques normales, en l'absence de cancer somatique ou d'effraction de la barrière hémato-testiculaire, les produits de ces gènes restent inconnus du système immunitaire. En revanche, lorsqu'ils sont exprimés de manière aberrante dans les cancers somatiques, ils peuvent être reconnus par le système immunitaire qui les considère comme des antigènes étrangers, induisant une réaction cellulaire et/ou humorale contre les cellules tumorales. Ainsi, du fait de leur expression tissulaire restreinte, les gènes CT et leurs produits constituent non seulement une source de marqueurs pour la détection des cancers, mais aussi des cibles idéales pour une immunothérapie anticancéreuse. Ces approches vaccinales des CT ont culminé avec le lancement en 2007 de tests cliniques de phase III sur un vaccin *MAGE-A3* pour les mélanomes et les cancers pulmonaires (GlaxoSmithKline GSK) (Brichard, 2008). Le but est de recruter et de suivre sur 10 ans 2 270 patients avec des tumeurs positives pour *MAGE-A3* auxquels seront administrés le vaccin ou un placebo. La première étude clinique publiée sur ce candidat, portant sur vingt-cinq patients souffrant de cancers pulmonaires au stade métastatique, avait permis d'obtenir sept réponses dont trois complètes, avec une excellente tolérance (Marchand et al., 1999), sans toutefois permettre d'identifier les mécanismes immunologiques mis en jeu dans cette réponse. Malgré ces avancées, relativement peu de progrès avaient été réalisés jusqu'à récemment dans la compréhension des fonctions des facteurs CT, que ce soit dans les cellules germinales ou dans les cellules cancéreuses, et des raisons de leur ré-expression aberrante dans les tumeurs.

3.3.2. Ré-expression des facteurs CT dans les cancers pulmonaires

Parmi la centaine d'antigènes CT, on distingue deux sous-groupes sur la base de leur localisation génomique (**Figure 30**) : une trentaine de gènes localisés sur le chromosome X (appelés gènes CT-X), et les autres. Les CT-X partagent un certain nombre de caractéristiques, dont la présence de multiples copies sur le génome issues d'événements de duplication (Ross et al., 2005). Dans le testicule, les gènes CT-X sont plutôt exprimés dans les spermatogonies en prolifération (Simpson et al., 2005). Les plus hautes fréquences de ré-expression des gènes CT-X sont observées dans les mélanomes, les cancers du poumon et de l'ovaire. A l'inverse, les gènes CT qui ne sont pas portés par le chromosome X sont le plus souvent présents en une copie unique, sont distribués à travers le génome et sont plutôt exprimés dans les stades plus tardifs de la spermatogenèse (Simpson et al., 2005). De nombreux gènes CT, dont *MAGE*, *BAGE*, *GAGE* et *NY-ESO-1* sont exprimés ectopiquement dans les carcinomes bronchiques, les principaux étant listés dans le tableau ci-contre (**Figure 30**). Parmi les cancers pulmonaires, ce sont les carcinomes épidermoïdes bronchiques (CEB) qui exhibent la plus haute fréquence d'expression d'antigènes CT, avec 90% des patients souffrant de CEB exprimant au sein de leur tumeur au moins un gène CT parmi neuf (*NY-ESO*, *MAGE-A3*, *MAGE-A4*, *MAGE-A10*, *CT7/MAGE-C1*, *SSX* et *SSX4*) (Chiriva-Internati et al., 2012).

Bien que certains îlots CpG soient hyperméthylés, le génome des cellules cancéreuses se caractérise par une hypométhylation génomique globale (voir paragraphe 3.1.2.1.2). En parfaite opposition avec l'inactivation par hyperméthylation de gènes suppresseurs de tumeur discutée auparavant (voir paragraphe 3.2.1.2), l'hypométhylation dans les cellules cancéreuses pulmonaires peut toucher des îlots CpG méthylés dans les tissus normaux, et pourrait donc aboutir à la réactivation de certains gènes (**Figure 29**) (Szyf et al., 2004). Les gènes présentant une expression restreinte au testicule (gènes TS pour testis-specific) exhibent justement un statut épigénétique particulier, car nombre d'entre eux présentent des promoteurs riches en CpG hyperméthylés dans les cellules somatiques où leurs expressions sont verrouillées (De Smet et al., 1999 ; Rousseaux, 2013). On a ainsi attribué l'expression ectopique des gènes *MAGE*, observée dans 70 à 85% des CBNPC, à la perte de méthylation des promoteurs (Jang et al., 2001). Par ailleurs, cette hypothèse selon laquelle l'expression aberrante des gènes CT dans les tumeurs serait due à une hypométhylation des promoteurs est confortée expérimentalement par les résultats qui démontrent une réactivation transcriptionnelle des gènes *GAGE* dans des cellules en culture traitées par l'agent déméthylant azadéoxycytidine (De Backer et al., 1999). Par conséquent, les gènes CT permettent de mieux appréhender certaines anomalies épigénétiques associées

à la transformation cellulaire et au développement des cancers. La perte dans les cellules somatiques de la répression des gènes CT peut être considérée comme le reflet direct de l'occurrence d'une méthylation anormale et donc comme un bon détecteur des anomalies épigénétiques se produisant dans ces cellules. Cela peut être exploité pour le développement de marqueurs indiquant la présence et l'évolution d'un processus tumoral, basés sur la détection des transcrits CT (Axelsen et al., 2007), des protéines ou encore des anticorps circulants dirigés contre ces CT (Chen et al., 1997).

La méthylation du promoteur ne semble pourtant pas être le seul mécanisme épigénétique garantissant la mise sous silence des gènes CT dans les cellules somatiques et permettant leur ré-expression dans les cellules cancéreuses. En effet, une acétylation des histones induite par le traitement de cellules somatiques par un inhibiteur des HDAC, la trichostatine A (TSA), permet une ré-expression du gène CT *GAGE* (D'Alessio et al., 2007). On a également montré que la déméthylation des régions promotrices de gènes CT (obtenue par une inhibition des DNMT) et la ré-expression qui en résulte sont associées à une modification du code histone sur ces gènes en faveur d'un gain de marques actives, comme une augmentation de l'acétylation de K9 et de la méthylation de K4 de l'histone H3 (James et al., 2006). On a suggéré alors que la déméthylation des CT se produirait en premier lieu et serait stabilisée secondairement par des modifications des marques d'histones, les protégeant contre une re-méthylation du promoteur et le rétablissement de la répression (De Smet et al., 2004 ; Lorient et al., 2006).

Néanmoins, une des particularités des antigènes CT-X est leur réactivation simultanée dans les cancers. Dans des études d'expression de neuf antigènes CT, il s'est avéré que 65% des mélanomes sont positifs pour trois ou plus de ces protéines et que 57% des tumeurs de poumon en expriment au moins deux (Sahin et al., 1998 ; Tajima et al., 2003). Par exemple, le gène *NY-ESO-1* est rarement exprimé en l'absence de *MAGEA3*. Ceci tend à remettre en cause la vision selon laquelle l'expression illégitime des gènes CT résulte d'une levée aléatoire de l'hyperméthylation de l'ADN, qui ne pourrait expliquer la survenue de ces co-expressions. Ainsi, les cellules tumorales subissent une hypométhylation globale de l'ADN, mais les mécanismes impliquant la levée spécifique de la répression des gènes CT reste incertains. Quelques indices sur ces mécanismes ciblés sont venus de la caractérisation du gène *BORIS*. La protéine BORIS est spécifiquement exprimée dans les cellules germinales mâles, où elle active l'expression d'autres gènes spécifiques du testicule (dont les gènes *CST* et *NY-ESO1*) en se fixant sur leurs promoteurs. L'expression illégitime de *BORIS* a été détectée dans un nombre significatif de cancers (D'Arcy et al., 2006 ; D'Arcy et al., 2008 ; Kholmanskikh et al., 2008 ; Risinger et al., 2007), dont les cancers pulmonaires. De manière intéressante, l'activation de BORIS dans les cellules somatiques induit l'expression d'autres facteurs CT, dont les gènes *MAGE-A* (Vatolin et al., 2005) et *NY-ESO1* (Hong et al., 2005). Dans des cellules pulmonaires en culture, on a démontré que BORIS se fixe aux promoteurs de *MAGE-A* et induit un relâchement de la chromatine (par déméthylation de l'ADN et/ou modification des marques d'histones) (Bhan et al., 2011) et que l'induction de l'expression de *NY-ESO1* par BORIS passe par le recrutement du facteur de transcription SP1 au promoteur (Kang et al., 2007). Ainsi, on peut envisager une première phase d'expression anormale de BORIS dans la cellule cancéreuse, dont l'expression illégitime entraînerait dans un second temps l'expression de nombreux autres gènes CT et la réactivation d'un programme germinale au sein de cellules somatiques.

3.3.3. Des facteurs CT comme moteurs de la carcinogenèse pulmonaire

Depuis longtemps, les similarités entre le développement des cellules germinales et le développement des tumeurs ont interpellé les chercheurs. Cellules germinales et cellules cancéreuses possèdent des caractéristiques phénotypiques communes, comme l'hypométhylation globale de l'ADN ou l'évasion immunitaire. Ces cellules présentent également des similarités troublantes comme l'immortalisation des

	Antigène CT	Identifiant CT	Fonction	Expression du gène	Détection de la protéine
Gènes CT-X	CT45A1	CT45	Inconnue	ADC (25%), CEB (48%)	CB (13%)
	DDX53 (CAGE)	CT26	Expression est dépendante du cycle : rôle dans la croissance cellulaire ?	CB (100%)	
	GAGE-1	CT4	Inconnue		CB (16%)
	MAGE-A1	CT1		ADC (14%), CBGC (14%)	CB (33%)
	MAGE-B1	CT3	Activateurs des ubiquitine ligases RING E3	CBNPC (11%)	
	MAGE-C1	CT7		ADC (18%), CBGC (17%), CEB (23%) CBPC (25%)	
	NXF2	CT39	Export des ARNm du noyau au cytoplasme	CB (15%)	
	NY-ESO-1	CT6	Inconnue	CBNPC (10%) : CEB (27%), ADC (14%), CBGC (22%)	CEB (20%), ADC (16%)
	SAGE	CT14	Inconnue	CBNPC (22%)	
	SSX1	CT5		CBNPC (38%) CBPC (50%)	
	SSX2	CT5	Régulateurs transcriptionnels	CBNPC (12%) CBPC (0%)	
	SSX4	CT5		CBNPC (38%) CBPC (0%)	
	XAGE-1	CT12	Inconnue	ADC (40%), CEB (55%)	ADC (53%)
Gènes CT-non X	BAGE	CT2	Inconnue	CBNPC (20%) CBPC (50%)	
	BRDT	CT9	Interaction avec les histones hyperacétylées	ADC (50%), CEB (41%), CBGC (44%)	
	CABYR	CT88	Voie de signalisation du calcium	CB (36%)	CB (40%)
	CAGE1	CT95	Inconnue	CBNPC (100%)	
	CASC5	CT29		CBNPC (41%)	
	CDCA1	CT106	Alignement et ségrégation des chromosomes durant la mitose	ADC (50%), CEB (75%)	CBNPC (74%)
	COX6B2	CT59		CB (24%)	
	CCDC62	CT109	Interaction avec le réticulum endoplasmique	CB (26%)	
	CEP55	CT111	Rôle dans la cytokinèse	ADC (94%)	CB (59%)
	DPPA2	CT100	Inconnue	CB (30%) Co-exprimé avec MAGE-A1	CBNPC (25%)
	SP17	CT22	Protéine de support pour les enzymes de la glycolyse	CB	
	SYCP1 (SCP1)	CT8	Assemblage du complexe synaptonémal au cours de la méiose	CB (7%)	

Figure 30 : Activation ectopique de gènes de la lignée germinale mâle (gènes CT) dans les cellules cancéreuses pulmonaires

Ce tableau liste quelques uns des gènes CT activés de manière aberrante dans les cancers du poumon.

CB : carcinome bronchique ; CBPC : carcinome bronchique à petites cellules ; CBNPC : carcinome bronchique non à petites cellules ; ADC : adénocarcinome ; CEB : carcinome épidermoïde bronchique ; CBGC : carcinome bronchique à grandes cellules.

Adapté de (Chiriva-Internati et al., 2012)

cellules germinales souches et la transformation des cellules tumorales, ou encore comme la migration des cellules germinales primordiales et la dissémination des cellules métastatiques. En conséquence, la découverte des gènes CT a permis d'émettre une hypothèse intéressante : l'expression aberrante de gènes de la lignée germinale résulterait en une réactivation du programme gamétogénique normalement réprimé dans les cellules somatiques, ce qui pourrait entraîner les cellules vers la voie de la tumorigenèse (Old, 2001). Au cours des années passées, l'accumulation des connaissances sur les fonctions des facteurs CT dans les cellules cancéreuses a montré que ces protéines exprimées ectopiquement pouvaient jouer un rôle oncogène.

La famille MAGE est une famille majeure d'antigènes CT, comprenant 23 gènes CT-X, codant des protéines avec un grand domaine central nommé MAGE Homology Domain (MHD) qui permet les interactions protéine-protéine (Scanlan et al., 2002b). Globalement, les cellules cancéreuses pulmonaires exprimant les gènes MAGE sont plus résistantes à la mort cellulaire (Park et al., 2002). L'expression conjointe de MAGE-A2 et MAGE-A6 confère par exemple aux cellules cancéreuses en culture un avantage prolifératif et une plus grande chimio-résistance au paclitaxel (Duan et al., 2003). On a également montré que la suppression des expressions des MAGE dans des lignées cellulaires issues de mélanomes induit l'apoptose, un phénomène expliqué par l'action neutralisante des protéines MAGE sur p53 (Yang et al., 2007). En particulier, la protéine MAGE-A2 recrute l'histone déacétylase HDAC3 sur p53, s'opposant ainsi à l'acétylation et à l'activation de p53 à la suite d'un traitement génotoxique (Monte et al., 2006). Quant à la protéine MAGE-A11, elle inhibe l'enzyme PHD2 (prolyl hydroxylase 2), nécessaire au maintien d'un faible niveau de la protéine HIF (hypoxia-inducible factor). L'expression illégitime de MAGE-A11 pourrait alors contribuer à la progression du cancer, via l'activation de HIF, l'enclenchement de la réponse hypoxique, l'induction de l'expression du VEGF et la promotion de l'angiogenèse (Aprelikova et al., 2009).

D'autre part, la dérégulation de gènes normalement exprimés dans les cellules spermatogéniques pourrait avoir des conséquences dramatiques sur l'organisation génomique des cellules somatiques. En effet, l'une des réorganisations les plus spectaculaires du génome se produit lors de la spermatogenèse. Au cours de la phase méiotique, le génome subit des recombinaisons génétiques, puis pendant la phase post-méiotique les histones sont progressivement remplacées par les protamines, des protéines basiques permettant une compaction drastique du génome dans le spermatozoïde (Rousseaux et al., 2008). De ce fait, les cellules spermatogéniques expriment de manière spécifique un ensemble d'effecteurs génétiques et de régulateurs épigénétiques très puissants (Rousseaux et al., 2011). Il est donc raisonnable de penser que leur expression illégitime pourrait grandement contribuer à l'acquisition de l'instabilité génétique ou au bouleversement du paysage épigénétique des cellules somatiques et donc à la transformation cellulaire (Rousseaux and Khochbin, 2009). Ainsi, le gène CT *SCP* est un composant du complexe synaptonémal, structure multi-protéique impliquée dans l'appariement des chromosomes homologues et la recombinaison au cours de la méiose (Ollinger et al., 2005). On peut imaginer que l'expression ectopique dans les cellules cancéreuses de ce facteur méiotique pourrait conduire à des ségrégations chromosomiques aberrantes, à des aneuploïdies et pourrait favoriser l'apparition d'une instabilité génomique, base de la tumorigenèse. De la même manière, le gène CT *SPO11* code une protéine qui induit des coupures double brin de l'ADN lors des recombinaisons méiotiques (Baudat and de Massy, 2004). Exprimée dans les cellules cancéreuses, elle pourrait favoriser l'apparition de translocations chromosomiques.

Confirmant le fait que les facteurs CT puissent être moteurs de la transformation cellulaire et de la croissance tumorale, des études menées sur des cohortes de patients souffrants de cancers pulmonaires de types histologiques variés tendent à démontrer une association entre les expressions de gènes CT et un pronostic sombre. C'est le cas des expressions de MAGE et GAGE qui ont été corrélées à une plus faible survie des patients CBNPC, principalement dans les cas d'adénocarcinomes (Grah et al., 2008 ; Gure et al., 2005 ; Melloni et al., 2004 ; Shigematsu et al., 2010 ; Yanagawa et al., 2011a ; Yoshida et al., 2006). Les déméthylations de CT et leurs expressions ectopiques dans les carcinomes bronchiques se révèlent donc être de bons marqueurs de pronostic en plus de permettre des thérapies vaccinales anti-cancéreuses. Outre leur grand intérêt pour les

applications médicales, l'étude des CT représente une voie de recherche novatrice et puissante pour la compréhension des mécanismes fondamentaux régulant la programmation du génome au cours de la différenciation cellulaire.

3.3.4. Un concept étendu aux facteurs tissu-spécifiques

Lorsque l'on étudie les bases de données d'expression génique (données EST « *expressed sequence-tag* », données transcriptomiques...), on s'aperçoit que le testicule est l'organe qui exhibe le plus d'expressions géniques spécifiques, c'est-à-dire de gènes dont l'expression est restreinte exclusivement à cet organe et réprimée dans l'ensemble des autres tissus de l'organisme (Rousseaux, 2013). Ceci peut expliquer en partie pourquoi la plupart des expressions ectopiques décrites jusqu'alors sont des expressions de gènes spécifiques du testicule. Néanmoins, il est légitime de proposer que certains autres gènes tissu-spécifiques puissent être également ré-exprimés de manière aberrante dans les cellules cancéreuses dont l'épigénome est profondément bouleversé (Berdasco and Esteller, 2010). Ces activations aberrantes de gènes, peu étudiées jusqu'alors, représente une source prometteuse de biomarqueurs tumoraux.

Certains gènes CT (comme *MAGE*, *CABYR* ou *SSX*) peuvent être exprimés non seulement dans les cellules testiculaires mais également dans d'autres tissus somatiques, comme le cerveau ou le placenta. Les gènes spécifiques du placenta ré-exprimés dans les cellules cancéreuses pourraient, comme les gènes CT, constituer de bons marqueurs de tumeur, et même contribuer à l'acquisition du phénotype cancéreux. Cette hypothèse est particulièrement séduisante lorsque l'on prend en compte les similarités entre le processus naturel d'envahissement de la muqueuse utérine par les cellules du trophoblaste et l'invasivité des cellules cancéreuses lors du processus métastatique. Cette observation nous permet de proposer une extension au concept de gènes CT : tout gène présentant une spécificité tissulaire ou temporelle, normalement réprimé dans les cellules de la muqueuse bronchique, mais ré-exprimé de manière aberrante dans les cellules cancéreuses pulmonaires pourrait participer à la transformation cellulaire. La protéine ectopique verrait ainsi sa fonction participer, hors de son contexte cellulaire habituel, au processus tumoral.

L'une des expressions hors contexte les mieux décrites, et récompensée par le Prix Nobel de médecine en 2009 (Szostak, 2010), est celle de la télomérase. La télomérase est une enzyme exprimée lors des étapes précoces du développement et dans les cellules souches adultes, mais l'expression du gène *hTERT* codant sa sous-unité catalytique est réprimée dans les cellules différenciées (Collins and Mitchell, 2002). Cette enzyme assure une fonction de transcription inverse au niveau des séquences d'ADN télomériques, permettant de lutter contre le raccourcissement des télomères au cours des divisions cellulaires. Or cette érosion irrévocable des télomères lors de la réplication de l'ADN passe pour l'un des phénomènes qui limite la capacité proliférative des cellules, étant donné que lorsque qu'une longueur critique est atteinte, la cellule entre en sénescence (Hayflick, 1997). La télomérase est souvent ré-exprimée dans les cellules cancéreuses, leur permettant d'éviter la « crise des télomères » et d'acquérir une capacité proliférative théoriquement illimitée (Counter et al., 1992). La télomérase est présente dans 90% des tumeurs en général, dans 80% des CBNPC et près de 100% des CBPC (Lantuejoul et al., 2007 ; Lantuejoul et al., 2004). Les tumeurs restantes semblent utiliser une voie alternative pour limiter l'érosion de leurs télomères appelée Alternative Lengthening of Telomeres (ALT) et basée sur la recombinaison homologue (Cesare and Reddel, 2010). Les mécanismes moléculaires précis permettant la ré-expression de la télomérase dans les cellules cancéreuses pulmonaires restent à définir (Masuda and Takahashi, 2002) mais quelques candidats ont été proposés : certains membres de la famille Myc sont capables de transactiver le gène *hTERT* (Wang et al., 1998 ; Wu et al., 1999) et l'expression de la télomérase pourrait alors être la conséquence de la fréquente sur-expression de Myc dans les cancers (Bieche et al., 2000 ; Latil et al., 2000 ; Sagawa et al., 2001). De nombreuses études ont cherché à savoir si la longueur des télomères ou l'activité télomérase étaient corrélées au risque de développer un carcinome bronchique ou au pronostic d'une tumeur pulmonaire (Jang et al., 2008 ; Jeon et al., 2012). Cependant, les résultats se sont avérés contradictoires et il n'apparaît pas possible aujourd'hui de mettre en lumière une association claire entre télomérase et

susceptibilité, stade tumoral ou survie dans le cancer du poumon ([Gansner and Rosas, 2013](#)). Néanmoins, des inhibiteurs de télomérase sont actuellement en tests cliniques pour des patients CBNPC ([Ruden and Puri, 2013](#)).

Un autre exemple de réexpression ectopique d'un gène exprimé normalement au cours des premières étapes du développement a été fourni par l'étude d'un membre de la famille DNMT. La protéine DNMT3L, qui ne possède pas d'activité enzymatique propre, interagit avec et régule l'activité des méthyltransférases de l'ADN DNMT3a et DNMT3b. La protéine DNMT3L est exprimée au cours de l'embryogenèse précoce et dans les cellules germinales, et est normalement réprimée dans les cellules somatiques. Une déméthylation de la région promotrice du gène *DNMT3L* est détectée dans les cellules cancéreuses et associée à une expression aberrante du gène ([Gokul et al., 2007](#)). Du fait du rôle régulateur important de DNMT3L dans la régulation de méthylation *de novo* ([Ooi et al., 2007](#)), il est facilement envisageable qu'une activation illégitime de DNMT3L puisse avoir un impact non négligeable sur l'épigénome, ce qui induirait une méthylation erronée du génome, favorisant par exemple l'hyperméthylation de gènes suppresseurs de tumeur et donc la transformation maligne des cellules affectées.

Des événements de remaniements chromosomiques, liés à l'instabilité génétique des cellules cancéreuses pulmonaires, peuvent également aboutir à des expressions aberrantes mettant en jeu des facteurs tissu-spécifiques. C'est le cas d'une translocation courante dans les cancers pulmonaires qui conduit à la production d'une protéine de fusion à activité oncogène. Il s'agit de l'inversion paracentrique dans le petit bras du chromosome 2 conduisant au gène de fusion *EML4-ALK* ([Sasaki et al., 2010](#)). Cette fusion a été décrite en 2007 dans un sous-groupe de CBNPC (3 à 10% des cas), le plus souvent chez des patients non-fumeurs ([Soda et al., 2007](#)). Elle conduit à la production d'une protéine chimérique Echinoderm Microtubule-associated protein-Like 4 (EML4) - Anaplastic Lymphoma Kinase (ALK). Au contraire de la protéine EML4 qui est ubiquitaire, l'expression de la protéine ALK est normalement restreinte au système nerveux central ([Morris et al., 1997](#)). Sa fonction est peu à peu décrite et l'on sait à présent qu'il s'agit d'un récepteur membranaire à activité tyrosine kinase, qui passe pour jouer un rôle important dans le développement du cerveau ([Roskoski, 2013](#)). Dans les cellules cancéreuses pulmonaires, l'inversion n'ayant pas toujours lieu au même endroit, de nombreux variants de la protéine EML4-ALK ont été identifiés, mais le domaine tyrosine-kinase d'ALK et l'extrémité amino-terminale coil-coiled de EML4 sont toujours présents. Alors que la protéine ALK s'exprime normalement à la membrane des cellules nerveuses, la protéine de fusion localise dans le cytoplasme des cellules cancéreuses pulmonaires. L'oncogénicité de cette protéine de fusion a été démontrée *in vitro* et *in vivo* ([Soda et al., 2008](#)). Plus précisément, il semble que la partie EML4 autorise la dimérisation ligand-indépendante de la partie ALK, conduisant à l'activation constitutive de la kinase ALK. Cette activation activerait de multiples voies de signalisation (dont les cascades Ras/Mek/Erk et PI3K/Akt) qui contribuent à la survie et à la transformation cellulaire ([Shaw and Solomon, 2011](#)). A peine un an après la découverte de la translocation, des inhibiteurs de l'activité tyrosine kinase d'ALK ont été développés et s'avèrent prometteurs ([McDermott et al., 2008](#)). Ceux-ci conduisent à l'apoptose *in vitro* et permettent une régression tumorale dans un modèle de souris. Un an plus tard, des tests cliniques de phase I sur des patients EML4-ALK ont montré une réponse radiographique dans 60% des cas ([Kwak, 2009](#)). En 2011, le crizotinib, molécule inhibitrice d'ALK, a été autorisé sur le marché américain ([Pilotto et al., 2013](#)).

Ainsi, les perturbations génétiques et épigénétiques des cellules bronchiques peuvent aboutir à une reprogrammation des expressions géniques et à une crise identitaire, au cours de laquelle des gènes spécifiques d'autres types cellulaires sont exprimés de manière aberrante. Jusqu'à récemment, les activations ectopiques de gènes tissu-spécifiques normalement silencieux ont été relativement peu étudiées. Pourtant, au-delà d'un rôle évident en tant que biomarqueurs de tumeur, on peut proposer que certaines protéines ectopiques puissent participer à la transformation cellulaire. En conséquence, il apparaît prometteur de rechercher de manière systématique ces expressions hors-contexte et d'étudier leur éventuelle contribution au processus oncogénique.

Les travaux de notre équipe dans ce champ de recherche ont mené à la caractérisation d'un nouveau facteur CT, la protéine ATAD2 (CT137) ([Caron et al., 2010](#)). Cette protéine, à domaine AAA-ATPase et à bromodomaine, capable d'interagir avec la chromatine, est spécifique de la lignée germinale mâle et des cellules souches, mais la réactivation aberrante du gène a été observée dans de nombreux types de tumeurs. Cette expression ectopique est de plus associée à un pronostic sombre dans les cancers du poumon et du sein. Les travaux menés jusqu'alors démontrent un rôle d'ATAD2 dans l'apoptose induite par des drogues génotoxiques ([Caron et al., 2010](#)), mais également dans la prolifération cellulaire et la croissance tumorale en modèle murin (données non publiées). Ces effets pourraient être dûs à des modifications globales de la dynamique chromatinienne ainsi qu'au rôle d'ATAD2 en tant que coactivateur de Myc. Ces travaux, actuellement poursuivis au sein de notre équipe visent à décrypter les bases moléculaires du rôle oncogène d'ATAD2, et à proposer de potentielles stratégies thérapeutiques basées sur des inhibiteurs d'AAA-ATPase et de bromodomaine.

Un second travail collaboratif récent a porté sur la découverte de l'expression aberrante de la protéine nucléaire Cyclon dans les cancers immuns, impliquée dans l'acquisition d'une résistance au rituximab, un anticorps monoclonal chimérique dirigé contre la molécule de surface CD20 induisant la mort des cellules B et utilisé comme agent chimiothérapeutique des lymphomes non-hodgkiniens ([Emadali, 2013](#)).

Enfin, notre équipe s'est intéressée à l'activité oncogénique de la protéine de fusion BRD4-NUT dans les carcinomes de la ligne médiane (NUT midline carcinoma – NMT), des carcinomes peu différenciés et très agressifs. La translocation chromosomique t(15;19)(q13;p13) aboutit à la fusion de la protéine à double bromodomaine BRD4, orthologue de BRDT mais d'expression ubiquitaire, avec une protéine de fonction inconnue nommée NUT (NUclear protein in Testis), normalement exprimée uniquement dans les cellules spermatogéniques. Ces travaux ([Reynoird et al., 2010b](#)) ont mis en évidence des propriétés uniques pour cette protéine de fusion et un rôle oncogène direct. En effet, les bromodomains de BRD4 permettent l'association de la protéine de fusion aux zones acétylées de la chromatine, alors que NUT recrute l'acétyltransférase CBP/p300 et stimule son activité catalytique, créant des foyers de chromatine hyperacétylée. En conséquence, la majorité des protéines CBP/p300 cellulaires est séquestrée dans ces foyers, privant le reste du génome de ces enzymes essentielles et causant la perte de la réponse p53-dépendante. Cette étude a également démontré que la suppression des foyers de séquestration par un traitement avec la TSA entraîne le relargage de CBP/p300 et réactive la réponse apoptotique médiée par p53, ce qui fait des HDACi un traitement prometteur pour ces carcinomes agressifs ([Schwartz et al., 2011](#)).

Objectifs du travail de thèse

L'équipe de recherche « Epigénétique et signalisation cellulaire » dirigée par le Dr Saadi Khochbin s'intéresse aux mécanismes épigénétiques impliqués dans la programmation du génome, en particulier ceux dirigeant la reprogrammation du génome mâle au cours de la spermatogenèse. Les travaux de l'équipe se sont notamment focalisés sur le rôle des variants d'histones spécifiques du testicule (Govin et al., 2004 ; Govin et al., 2007 ; Montellier, 2013), des chaperonnes de protéines de transition (Govin et al., 2006), des modifications d'histones (Montellier et al., 2012 ; Tan et al., 2011) ainsi que de BRDT, un effecteur épigénétique puissant spécifique des cellules germinales mâles (Moriniere et al., 2009 ; Pivot-Pajot et al., 2003) impliqué dans la programmation et la réorganisation du génome mâle au cours des étapes méiotique et post-méiotique de la spermatogenèse (Gaucher et al., 2012).

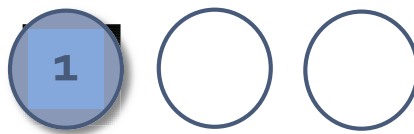
En parallèle, l'équipe a montré que plusieurs facteurs testiculaires impliqués dans la programmation du génome mâle, notamment BRDT, sont exprimés de manière aberrante dans des tumeurs affectant les tissus somatiques. En conséquence, un nouveau pan de recherche s'est ouvert au sein de l'équipe, basé sur l'hypothèse que ces expressions aberrantes dans les cellules cancéreuses de facteurs spécifiques du testicule puissant participer à la progression tumorale (Rousseaux and Khochbin, 2009). Les travaux de recherche translationnelle initiés alors se sont notamment appuyés sur une cohorte de patients souffrant des cancers pulmonaires, et dans les tumeurs desquels les expressions hors-contexte de gènes tissu-spécifiques ont été systématiquement recherchées et corrélées avec les données cliniques et anatomopathologiques (Rousseaux, 2013) (Annexe 2).

C'est dans ce contexte que j'ai initié mes travaux de doctorat, visant à étudier les conséquences de l'expression aberrante de certains de ces gènes tissu-spécifiques dans les cellules cancéreuses pulmonaires. Les données transcriptomiques et de survie des patients de la cohorte « Cancer du poumon » nous ont permis de sélectionner plusieurs gènes candidats qui, du fait de leur profil d'expression restreint, de leur ré-expression dans les cancers, de leur association avec un pronostic sombre et/ou du fait de leurs fonctions préalablement décrites dans leur contexte cellulaire normal, nous ont semblé être intéressants pour une étude plus approfondie de leur rôle dans les cellules cancéreuses. L'hypothèse sous-jacente était alors qu'il puissent servir de biomarqueurs de tumeur, de pronostic et/ou de réponse au traitement, voire de facteurs oncogènes ou de potentielles cibles thérapeutiques.

Mes travaux de thèse se sont principalement concentrés sur le gène **PRL**, codant l'hormone hypophysaire prolactine. Ces travaux ont découlé de l'identification du gène **PRL** dans la signature transcriptomique des tumeurs agressives des patients de la cohorte « Cancer du poumon ». Notre premier objectif a été de vérifier l'activation ectopique du gène **PRL** dans les cellules cancéreuses pulmonaires, à la fois chez les patients de la cohorte et dans des modèles de lignées cellulaires de cancer du poumon à petites cellules. Nous avons ensuite exploré les causes de la ré-expression de ce gène dans les tumeurs pulmonaires, ainsi que les conséquences de cette expression aberrante sur le phénotype cancéreux. Des analyses de survie nous ont permis de démontrer une association forte de l'expression du gène **PRL** avec un pronostic sombre, nous faisant envisager un rôle de l'activation de **PRL** dans l'agressivité des cellules tumorales. Dans des lignées cellulaires modèles, nous avons démontré que l'activation du gène **PRL** confère aux cellules cancéreuses pulmonaires une résistance accrue à l'apoptose induite par un stress. Enfin, lorsque nous avons voulu préciser les modes d'expression du gène **PRL** et d'action de la prolactine dans les cellules de cancer du poumon, nous avons mis à jour un mode d'action inédit, indépendant du récepteur à la prolactine. L'hypothèse d'une action *in cellulo* d'une prolactine non sécrétée que nous avons émise a été confortée par l'observation de la synthèse par les cellules cancéreuses pulmonaires d'un ARN messager **PRL** tronqué de ses premiers exons. Ces travaux, présentés dans la **partie principale** des résultats, font émerger la possibilité d'utiliser cette expression ectopique de **PRL** comme marqueur de pronostic et potentielle cible thérapeutique dans le traitement des tumeurs pulmonaires agressives.

Des travaux complémentaires, réalisés préalablement sur trois candidats spécifiques du testicule dont nous avons observé l'activation ectopique dans les cellules cancéreuses pulmonaires (**BRDT**, **SOX3o** et **SPATA22**), sont présentés dans les **deuxième et troisième parties** des résultats. L'ensemble des protocoles utilisés dans ce travail sont présentés dans la section « **Matériels et méthodes** » à la fin de ce manuscrit. Ce travail de thèse n'a pas encore donné lieu à la publication d'un article scientifique. Cependant, les bases théoriques justifiant notre intérêt pour les expressions ectopiques de gènes tissu-spécifiques dans les cancers ont été publiées dans une revue ([Wang et al., 2011](#)) présentée en annexe (**Annexe 1**).

RESULTATS



L'expression aberrante du gène hypophysaire *PRL*
dans les tumeurs pulmonaires
est associée à une chimio-résistance accrue et à un pronostic sombre :

Vers un rôle oncogène récepteur-indépendant d'une prolactine tronquée
non sécrétée

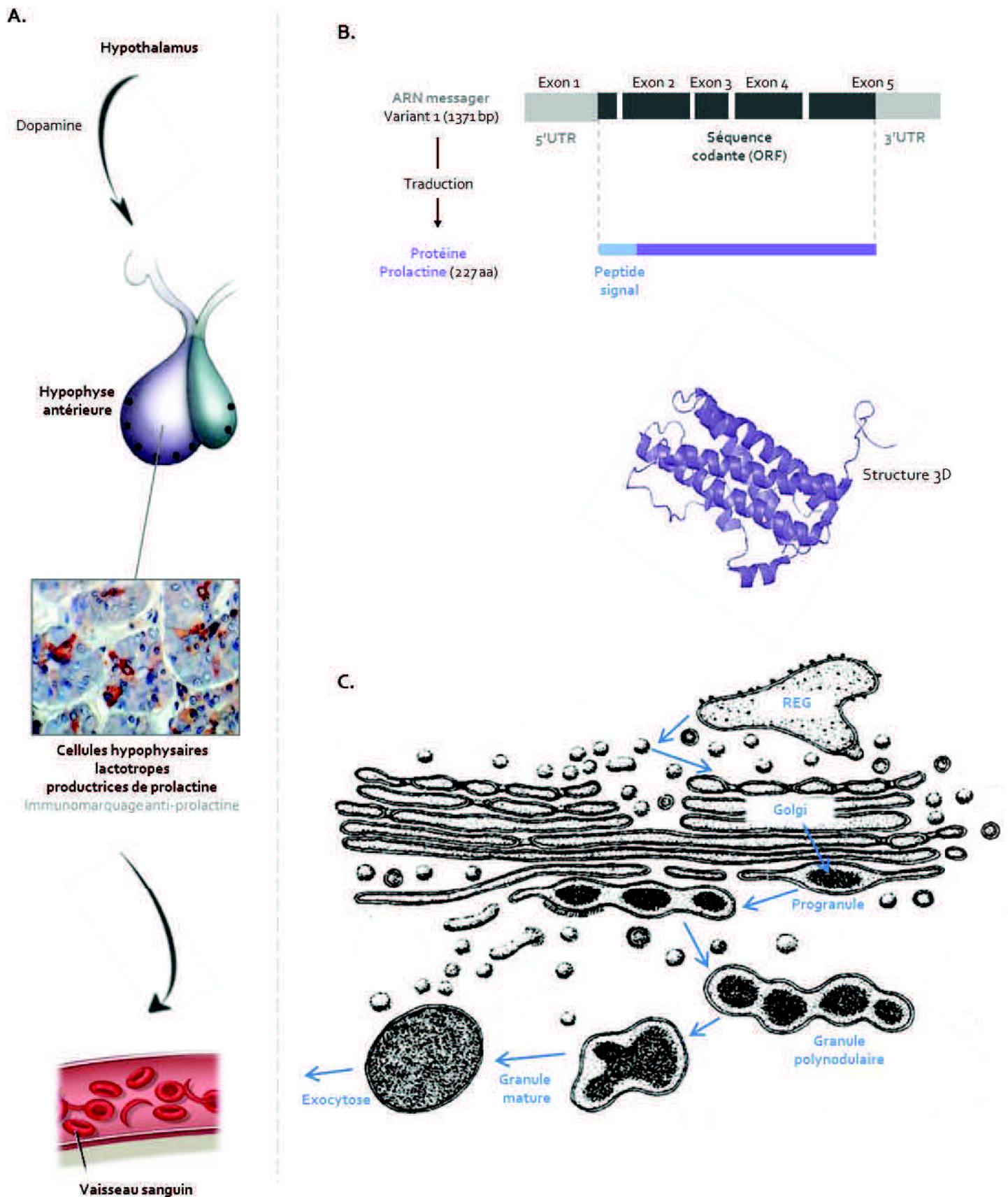


Figure 31 : Sécrétion de prolactine par les cellules lactotropes de l'anté-hypophyse

A. L'axe hypothalamo-hypophysaire lactotrope.

B. La traduction de l'ARN messager hypophysaire conduit à la production d'une protéine de 227 acides aminés et 25 kDa. Séquence 3D tirée de (Keeler et al., 2003).

C. Grâce à son peptide signal, la prolactine emprunte la voie de sécrétion via le réticulum endoplasmique granuleux (REG) et l'appareil de Golgi. La prolactine est stockée sous forme d'aggrégats protéiques dans des granules de sécrétion. L'exocytose provoque la libération de la prolactine dans la circulation sanguine.

Adapté de (Dannies, 2002).

1. Contexte de l'étude

1.1. La prolactine, hormone peptidique hypophysaire

1.1.1. Structure de la protéine

La protéine prolactine appartient à une famille d'hormones contenant l'hormone de croissance (Growth Hormone GH) et l'hormone lactogène placentaire (Placental lactogen PL), famille qui résulte probablement de la duplication d'un gène ancestral (Cooke et al., 1981 ; Niall et al., 1971). La traduction du cadre ouvert de lecture (Open Reading Frame ORF) de l'ARN messenger PRL, commun aux variants 1 hypophysaire et 2 extra-hypophysaire, aboutit à la synthèse de la pré-prolactine de 227 acides aminés d'une masse moléculaire de 25,88 kDa (Bole-Feysot et al., 1998) (Figure 31B). Cette protéine peut subir différentes modifications post-traductionnelles, incluant glycosylation, phosphorylation et clivage protéique (Walker, 1994). La prolactine mature ne contient plus que 199 acides aminés après clivage du peptide signal à l'extrémité amino-terminale. De plus, 6 résidus cystéines participent à l'établissement de 3 ponts disulfures intra-moléculaires (Cys 4-11, 58-174 et 191-199). La structure 3D de la prolactine a été décryptée il y a dix ans par résonance magnétique nucléaire (RMN) (Keeler et al., 2003). La PRL contient un cœur de 4 hélices alpha organisées en configuration up-up-down-down (Figure 31B). La structure tridimensionnelle diffère légèrement de celle de la GH, notamment du fait de la présence d'une boucle amino-terminale structurée en contact avec la première hélice. Les résidus nécessaires à l'interaction de la PRL avec son récepteur sont regroupés en surface.

1.1.2. Sécrétion par les cellules hypophysaires

La prolactine est une hormone hypophysaire, produite par les cellules lactotrophes qui constituent 20 à 50% des cellules du lobe antérieur de l'hypophyse (Figure 31A) (Ignacak et al., 2012). L'existence de la prolactine a été découverte au début du vingtième siècle, suite à l'observation des changements histologiques de l'hypophyse chez les femmes enceintes (Riddle, 1933 ; Stricker, 1928). Grâce à son peptide signal, la prolactine produite par les cellules lactotrophes emprunte la voie de sécrétion via le réticulum endoplasmique et l'appareil de Golgi (Figure 31C) (Keeler et al., 2003). Elle est ensuite stockée sous forme d'aggrégats protéiques dans des granules de sécrétion (Dannies, 2002 ; Dannies, 2012) puis elle est sécrétée et libérée dans le flot sanguin (Figure 31A). La prolactinémie est habituellement comprise entre 5 et 20 µg/L, chez les hommes comme chez les femmes (Mancini et al., 2008). La sécrétion de prolactine étant pulsatile, sa concentration sérique présente des variations, notamment circadiennes (concentration plus haute en phase de sommeil) mais aussi selon le cycle menstruel chez la femme (Parker et al., 1973). Durant la grossesse, les œstrogènes stimulent la prolifération des cellules lactotrophes et l'hypophyse double de taille. La prolactine est alors abondamment produite et les niveaux sériques dépassent 100 µg/L.

La sécrétion de prolactine par les cellules hypophysaires est sous le contrôle des cellules neuroendocrines de l'hypothalamus (Figure 31A). Dès les années 1960, on avait observé qu'un facteur hypothalamique PIF (Prolactin Inhibiting Factor) s'avérait capable d'inhiber la libération de la prolactine (Schally et al., 1965) et on a finalement identifié ce facteur comme étant la dopamine (Clemens and Shaar, 1980). La dopamine, produite par les neurones TIDA (tuberoinfundibular dopaminergic neurones) du noyau arqué, est libérée via le système porte dans les capillaires sinusoides de l'anté-hypophyse, agit via les récepteurs à la dopamine D₂ de la membrane des cellules lactotrophes et inhibe la sécrétion de prolactine. En retour, la prolactine apparait capable de moduler la sécrétion de dopamine par les neurones TIDA, mettant en lumière l'existence d'un rétrocontrôle dans cet axe hypothalamo-hypophysaire lactotrope (Lyons et al., 2012). A l'inverse, de nombreux facteurs semblent capables de stimuler la libération de prolactine, dont l'hormone thyroïdienne (Thyrotropine Releasing Hormone TRH), la ghréline ou l'ocytocine (Messini et al., 2010).

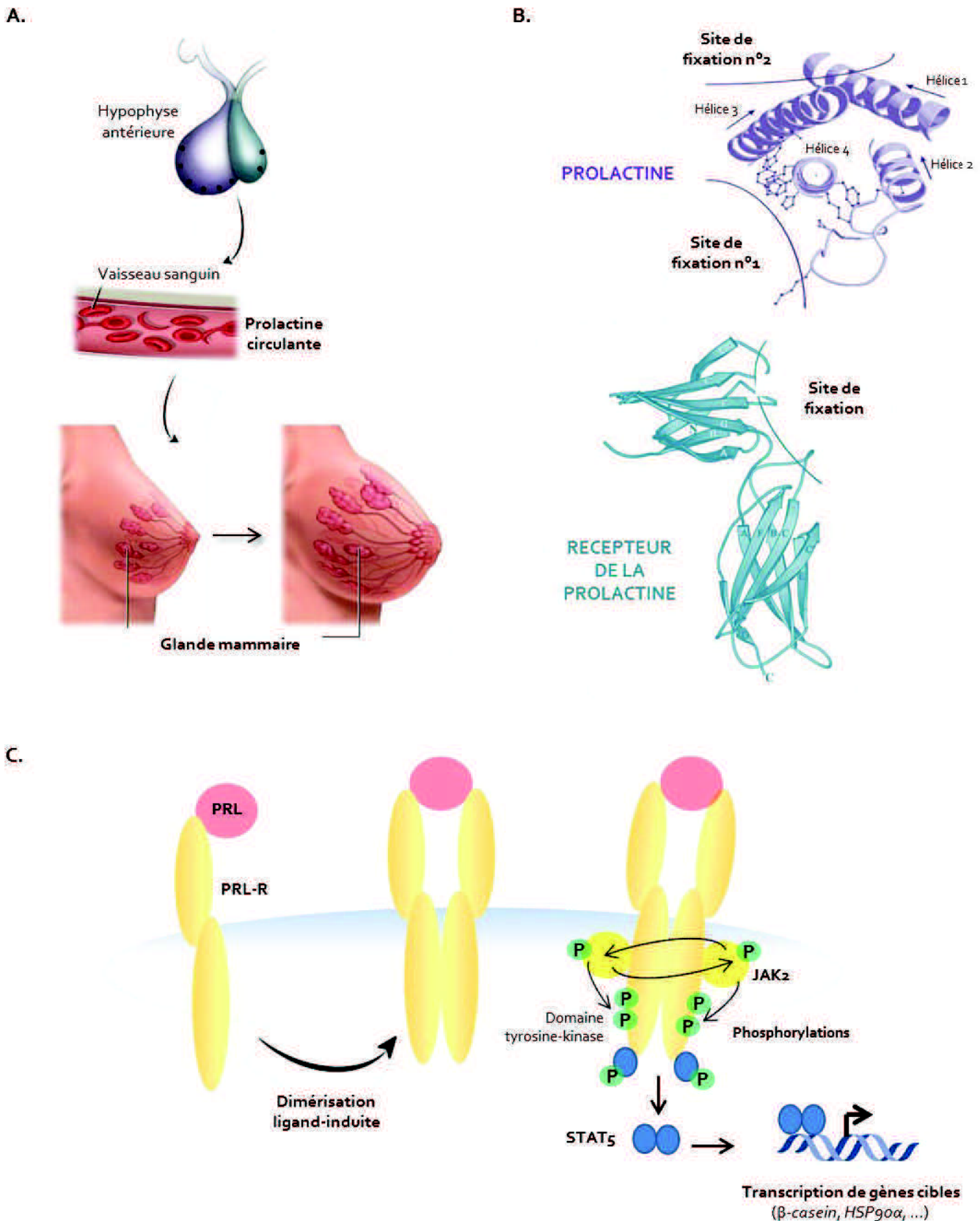


Figure 32 : Le récepteur PRL-R médie l'effet mammotrope et lactogénique de la prolactine

A. La prolactine circulante agit sur son organe cible, les glandes mammaires, pour promouvoir la lactation.

B. La prolactine possède deux sites de fixation et assure la dimérisation ligand-induite du récepteur.

Adapté de (Bole-Feyssot et al., 1998).

C. Suite à la liaison de la prolactine au domaine extra-cellulaire, le récepteur se dimérise et les domaines tyrosine-kinase sont activés, causant l'autophosphorylation des domaines TK et de la protéine JAK-2. La protéine STAT5, phosphorylée, se dimérise et est transloquée dans le noyau où elle agit comme facteur de transcription, promouvant l'expression des gènes nécessaires à la formation du lait maternel.

Adapté de (Ignacak et al., 2012).

1.2. Le rôle lactogénique et mammotrope de la prolactine

La prolactine est une hormone qui agit selon un fonctionnement endocrine classique : libérée dans le sang par l'hypophyse, elle emprunte la circulation générale et agit sur son organe cible principal, le sein, qui possède des récepteurs spécifiques (**Figure 32A**). La plupart des enseignements sur les fonctions de la prolactine sont venus de modèles animaux. On a ainsi montré que la prolactine était capable d'induire la production de lait chez le lapin ([Stricker, 1928](#)). De fait, l'hormone hypophysaire exerce un double effet sur les glandes mammaires : un effet mammotrope (croissance des glandes mammaires) et un effet lactogénique (stimulation de la synthèse du lait). L'effet mammotrope se manifeste dès la puberté sous l'influence hormonale (prolactine, mais aussi stéroïdes ovariens, hormone de croissance et glucocorticoïdes) avec le développement des canaux galactophores et des bourgeons glandulaires ([Sternlicht, 2006](#)). Puis l'augmentation de la prolactinémie au cours d'une grossesse - en conjonction avec l'oestradiol, l'hormone lactogène placentaire, l'insuline et le cortisol - cause l'allongement des canaux galactophores, ainsi que la multiplication et le développement des acini (unités ducto-lobulaires de production du lait) ([Briskin et al., 1999](#)). La production de lait ne commence quant à elle qu'après la naissance, lorsque les niveaux d'œstrogènes et de progestérone s'effondrent ([Rosen et al., 1978](#)) et que le stimulus de tétée est présent ([de Greef et al., 1987](#)). Sous l'effet de la prolactine, la pression sanguine augmente dans les capillaires entourant les acini, permettant le passage vers les cellules épithéliales mammaires des éléments nécessaires à la fabrication du lait. Dans les cellules de l'acinus, il y a augmentation de la synthèse des enzymes de biosynthèse des lipides ([Rudolph et al., 2007](#)) et des protéines composant le lait (bêta-caséine notamment) ([Rudolph et al., 2003](#)). Le lait est ensuite déversé dans la lumière de l'acinus et transite par les canaux galactophores jusqu'au mamelon.

1.2.1. Le récepteur de la prolactine

Le récepteur à la prolactine (PRL-R) a été identifié dans les années 1970 ([Boutin et al., 1988](#) ; [Posner et al., 1974](#)). Le PRL-R fait partie de la superfamille des récepteurs à cytokines de classe 1 qui inclue également les récepteurs de la GH, des interleukines, du GM-CSF (granulocyte macrophage-colony stimulating factor), du LIF (leukemia inhibitory factor) ou de l'EPO (érythropoïétine) ([Bazan, 1990](#) ; [Kelly et al., 1991](#)). Le récepteur à la prolactine contient une région extracellulaire (extra-cellular domain ECD) de 210 acides aminés, capable de lier la prolactine (**Figure 32B**), comprenant deux sous-domaines D1 et D2 à feuillets bêta analogues au module fibronectine de type III ([Bazan, 1990](#)). La protéine comprend également une région transmembranaire de 24 acides aminés, et une région cytoplasmique (intracellular domain ICD) contenant un domaine BOX-1 riche en prolines, connu pour être un site d'ancrage de la kinase JAK-2. Exprimé par les cellules des glandes mammaires, le PRL-R va être responsable des effets de la prolactine sérique sur cette organe, via une voie de signalisation en aval ([Clevenger and Kline, 2001](#)).

1.2.2. La voie de signalisation lactogénique

La fixation de la prolactine sur le PRL-R induit la dimérisation du récepteur. En effet, l'hormone présente deux sites de liaison distincts, capables de lier les domaines extracellulaires de deux récepteurs membranaires (**Figure 32B**) ([Broutin et al., 2010](#)). Ceci provoque un changement conformationnel de ce complexe ternaire et l'activation des récepteurs via le rapprochement physique des domaines cytoplasmiques à activité tyrosine-kinase (**Figure 32C**) ([Clevenger and Kline, 2001](#)). La phosphorylation de la kinase JAK-2 (Janus kinase 2), constitutivement associée au récepteur sur la face cytoplasmique ([Pellegrini and Dusanter-Fourt, 1997](#)), est l'évènement qui initie l'activation de la voie JAK-STAT ([Watson and Burdon, 1996](#)). Les protéines STAT, particulièrement le membre STAT5, sont les effecteurs majeurs de la transduction du signal de la prolactine. Le facteur STAT5, appelé initialement MGF (pour Mammary Gland Factor) ([Schmitt-Ney et al., 1991](#) ; [Wakao et al., 1994](#)) s'associe au complexe récepteur-ligand, est phosphorylé, se dimérise puis rejoint le noyau (**Figure 32C**). Suite à sa fixation sur l'ADN sur une séquence GAS (gamma-interferon-activated sequence), le dimère STAT

active la transcription de gènes cibles (Gouilleux et al., 1994 ; Ihle, 1996), dont les gènes codant la bêta-caséine, la lactoglobuline ou encore la protéine de lactosérum acide – toutes constituants du lait maternel (Groner and Gouilleux, 1995 ; Guyette et al., 1979). L'action transactivatrice de STAT5 apparaît médiée par une interaction avec le coactivateur CBP/p300 (Pfitzner et al., 1998).

Au-delà de cette action sur l'expression des protéines du lait par l'activation de la voie JAK/STAT, la prolactine est aussi plus largement impliquée dans la prolifération, la différenciation et la survie des cellules de l'épithélium mammaire. Le facteur STAT5 est capable de se lier aux séquences GAS présentes dans le promoteur du gène codant la cycline D1 (Brockman et al., 2002) régulant le cycle cellulaire. Les actions mitogènes et morphogènes de la prolactine semblent également impliquer d'autres voies de signalisation intracellulaires en aval du récepteur PRL-R, dont la voie des MAP kinases GRB2/SBC/SOS/Ras/RAF/MAPK et la voie de la Src kinase (Fresno Vara et al., 2001). Ces cascades de signalisation contrôlent l'activité de gènes impliqués dans la prolifération et la croissance de l'épithélium mammaire.

1.3. D'autres cibles et d'autres sources de la prolactine

1.3.1. Les cibles secondaires de l'hormone hypophysaire

Alors que la glande mammaire était considérée comme la cible principale de la prolactine, on a peu à peu observé de nouvelles et nombreuses localisations du récepteur à la prolactine. L'expression du gène *PRL-R* a été décrite dans de nombreux tissus en dehors de l'épithélium mammaire, tels que l'utérus, le rein, le foie, la prostate, l'ovaire, la peau et certaines aires du système nerveux central, les cellules immunes... (Chavez-Rueda et al., 2005 ; Hair et al., 2002). Ceci peut théoriquement conférer à la prolactine circulante la capacité de moduler un large panel de processus physiologiques. De fait, on considère aujourd'hui que la prolactine est impliquée dans plus de 300 fonctions au sein de l'organisme, comme la reproduction, le métabolisme, la fonction immune, l'osmorégulation ou l'ostéogenèse (Freeman et al., 2000 ; Goffin et al., 2002 ; Grattan and Kokay, 2008 ; Hair et al., 2002 ; Ignacak et al., 2012). Il faut tout de même noter que la compréhension des multiples fonctions de la prolactine au sein du corps humain reste limitée du fait de l'absence de maladie génétique humaine liée aux gènes *PRL* ou *PRL-R*. Sans modèle clinique de déficience isolée de la signalisation prolactine, l'identification des fonctions qui dépendent ou sont modulées par la prolactine reste partielle et n'a souvent reposé que sur l'étude des modèles animaux (Bernichtein et al., 2010). La compréhension de la signalisation prolactine est également compliquée par l'existence de plusieurs isoformes du récepteur de la prolactine, dont les plus communes sont les formes longue (LF), intermédiaire (IF) et courte (SF) (Horseman, 2002). Ces isoformes découlent d'épissages alternatifs des ARNm *PRL-R*, mais aussi de clivages protéolytiques (Ignacak, 2012). Ces isoformes sont exprimées de manière variable au sein des tissus mentionnés auparavant et semblent présenter des différences fonctionnelles (Binart et al., 2010).

Enfin, on a mis en évidence un mode d'action original de la prolactine circulante lorsqu'elle est tronquée par des protéases. Le clivage protéique de la prolactine dans la région amino-terminale aboutit à la formation de différentes formes courtes, dont la forme de 16 kDa. Ces formes, renommées vasoinhibines, semblent jouer un rôle sur les cellules endothéliales. On a décrit des actions anti-proliférative, anti-angiogénique, anti-migratoire et pro-apoptotique du variant 16 kDa sur ces cellules (Clapp et al., 1993 ; Lee et al., 2007 ; Tabruyn et al., 2003), ce qui conduirait à la vasoconstriction, à la perte de perméabilité vasculaire et à l'inhibition de la formation de nouveaux vaisseaux (Clapp et al., 2006 ; D'Angelo et al., 1999). Ce clivage protéolytique de la prolactine en peptides vasoinhibines a été impliqué dans diverses pathologies, dont la rétinopathie liée au diabète (Triebe et al., 2011) ou les cardiomyopathies post-partum (Hilfiker-Kleiner et al., 2007). Néanmoins, les modes de clivage et d'action de la prolactine 16 kDa restent obscurs. Plusieurs enzymes à activité protéolytiques ont été suggérées (dont les métalloprotéases MMP de la matrice extracellulaire et la cathepsine D, lysosomale ou sécrétée) et pourraient agir spécifiquement dans certains tissus (rétine, myocarde), à l'intérieur ou hors des

cellules (Macotella et al., 2006 ; Piwnica et al., 2006). Cependant les modes de régulation du clivage restent discutés et aucun récepteur spécifique de cette forme courte n'a été identifié (Lkhider et al., 2010).

1.3.2. Les sources extra-hypophysaires de prolactine

Alors que l'on s'est longtemps focalisé sur le rôle de la prolactine circulante hypophysaire, on a récemment pris en compte l'existence de sites extra-hypophysaires de production de la prolactine. Cette production locale a été documentée dans le sein (Fields et al., 1993) mais aussi dans les cellules déciduales (cellules de l'endomètre participant à la formation du placenta), la prostate, les lymphocytes, les cellules adipeuses ou les glandes lacrymales (Ben-Jonathan et al., 1996). Alors que dans l'hypophyse, l'expression du gène *PRL* est contrôlée par le facteur de transcription hypophysaire Pit-1 et repose sur l'activation du promoteur proximal (**Voir Introduction, paragraphe 1.1.1.3**), la production extra-hypophysaire de prolactine est décrite comme Pit-1-indépendante et régulée par le promoteur distal (Gerlo et al., 2006). En conséquence, cette sécrétion locale est insensible à tous les agents chimiques inhibant la sécrétion de la prolactine par l'hypophyse. On pense que cette prolactine extra-hypophysaire pourrait agir localement par des voies paracrine (sur une cellule adjacente) et/ou autocrine (sur la cellule sécrétrice elle-même). Ainsi, la sécrétion de prolactine par les lymphocytes T activés permet d'induire et d'entretenir la prolifération des cellules voisines et de la cellule sécrétrice, et participe à la production des cytokines. En effet, la liaison de la prolactine à son récepteur PRL-R exprimé à la surface des cellules immunes active la voie JAK-STAT qui stimule l'expression du gène *T-bet*, codant un facteur de transcription clé pour la production de l'interféron (Chavez-Rueda et al., 2005 ; Tomio et al., 2008)

Etant donné le rôle de la prolactine dans la prolifération et la survie des cellules mammaires, la question de l'implication de la prolactine dans la tumorigenèse humaine, cancer du sein en particulier, s'est rapidement posée, a longtemps été débattue et reste un champ de recherche très actif (Fernandez et al., 2010). Au cours des trente dernières années ont été démontrés un rôle de la prolactine circulante sur le développement de tumeurs mammaires, mais également un rôle de la prolactine extra-hypophysaire, produite localement au sein des tumeurs du sein ou de la prostate, sur la transformation cellulaire (**voir Discussion**). Dans l'étude qui va suivre, nous avons observé pour la première fois une expression aberrante du gène *PRL* dans les cancers du poumon. Nos efforts pour préciser les conséquences de cette ré-expression hors-contexte sur le phénotype des cellules cancéreuses nous ont amenés à mettre en lumière l'implication de l'expression de *PRL* dans l'agressivité des cellules cancéreuses pulmonaires, un effet médié par des mécanismes inédits, totalement différents des modes d'action de la prolactine décrits dans cette introduction.

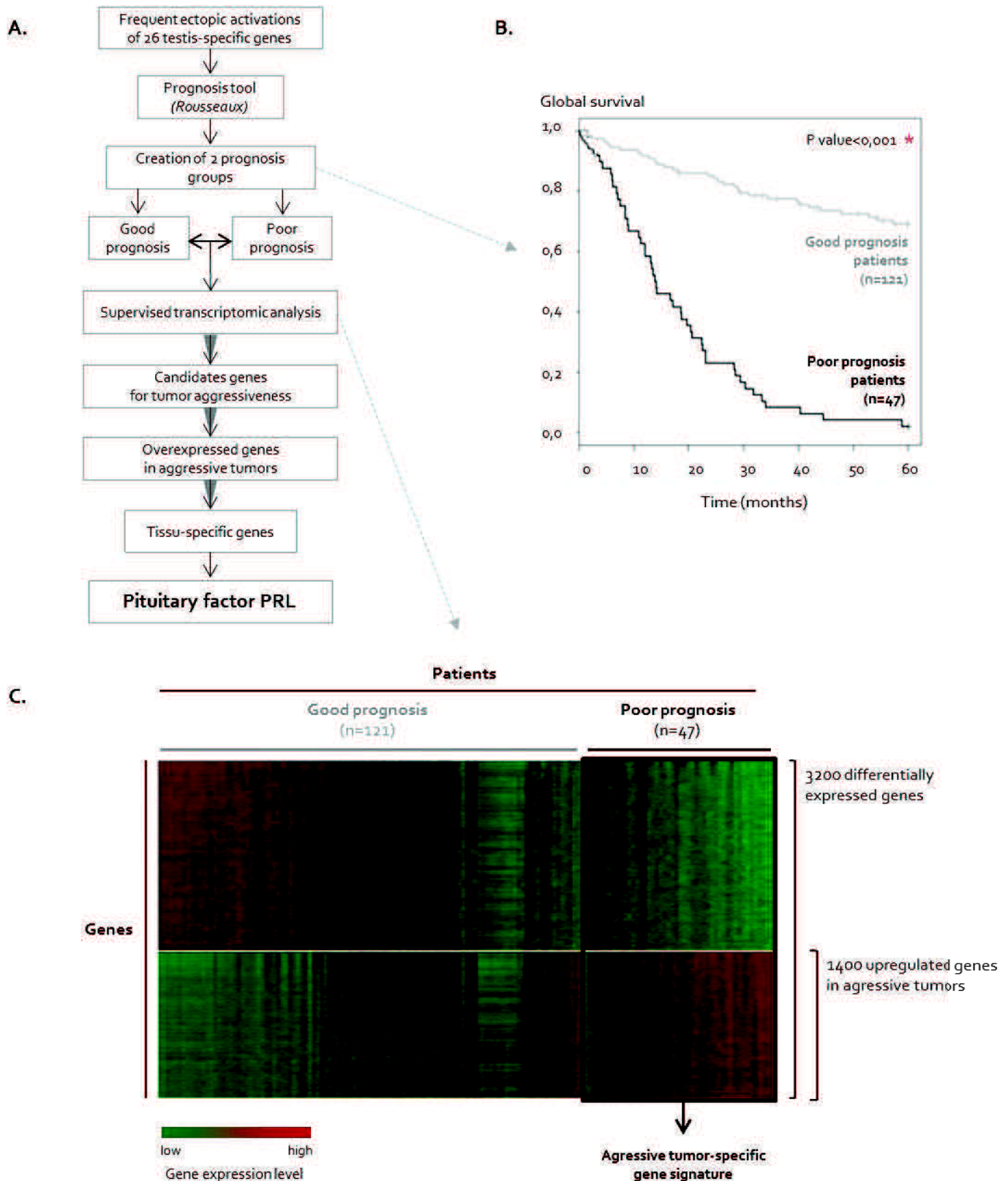


Figure 33: *PRL* gene is identified in the gene signature of aggressive tumors (Rousseaux, 2013)

A. Diagramm of the followed approach, aiming to identify expressions of tissue-specific genes in aggressive tumors of poor prognosis patients.

B. Survival probability curves (estimated by the Kaplan-Meier method) for good-prognosis (grey) and poor-prognosis (black) patients. The significance of the difference in prognosis is assessed by the univariate logrank test. * p value < 5%

C. Supervised transcriptomic analysis of good- versus poor-prognosis patients, based on the transcriptomic data of lung tumors obtained in Affymetrix technology. Expression levels within tumors are represented by a heat-map, for the selected 168 patients (horizontal) and for the 3200 genes identified as differentially expressed between good- and poor-prognosis patients (vertical). These genes have been clustered, allowing to discriminate a subset of 1400 genes upregulated in aggressive tumors (bottom right hand corner). *PRL* gene belongs to this upregulated subset of genes.

2. Résultats

2.1. Le gène hypophysaire *PRL* est exprimé de manière aberrante par les cellules cancéreuses pulmonaires

2.1.1. Le gène *PRL* est identifié dans la signature transcriptomique de tumeurs pulmonaires agressives

Je vais ici illustrer succinctement les étapes nous ayant conduit à sélectionner le gène *PRL* comme candidat associé à l'agressivité des tumeurs pulmonaires (**Figure 33**). Notre intérêt pour le gène *PRL* s'est manifesté suite à une étude menée dans notre équipe et récemment publiée ([Rousseaux, 2013](#)) (**Annexe 2**). Ce travail s'est appuyé sur une cohorte de 293 cancers pulmonaires (patients de la cohorte « Cancer du poumon » établie au CHU de Grenoble par les Professeurs Elisabeth et Christian Brambilla (**Voir Matériels**)) et sur les données transcriptomiques obtenues sur ces tumeurs (étude GSE30219 réalisée grâce à la technologie Affymetrix (**Voir Méthodes**)). Par l'analyse de ces données, Sophie Rousseaux a mis au point un outil pronostic permettant de distinguer deux sous-ensembles de patients (**Figure 33A**). Cet outil pronostic repose sur une combinatoire d'expressions de 26 gènes testiculaires fréquemment exprimés de manière aberrante dans les cancers pulmonaires. On isole les 121 patients qui n'expriment aucun de ces gènes au sein de leur tumeur d'une part, et les 47 patients qui en expriment trois ou plus d'autre part. On constate alors que ces deux groupes de patients présentent des survies très différentes (**Figure 33B**). Les 121 patients présentent un pronostic favorable : 70% d'entre eux sont encore en vie au terme de la période de suivi de 60 mois. Ils sont appelés en conséquence « Patients à bon pronostic ». A l'inverse, les 47 patients dont la tumeur pulmonaire exhibe une expression d'au moins 3 de ces gènes testiculaires ont un pronostic sombre et une survie significativement diminuée ($p\text{-value} < 0,001$). Ils sont renommés « Patients à pronostic sombre ». Cet outil pronostic nous a donc donné la capacité d'isoler spécifiquement les tumeurs pulmonaires agressives, causes de fréquentes rechutes et/ou métastases, et généralement d'une issue fatale à court terme ([Rousseaux, 2013](#)).

Une étude transcriptomique supervisée nous a ensuite permis de comparer les expressions géniques entre les tumeurs de bon pronostic et les tumeurs de pronostic sombre (**Figure 33C**). Cette analyse met en lumière des signatures d'expressions géniques distinctes entre les deux types de tumeurs avec 3200 gènes différentiellement exprimés entre ces deux groupes. De manière intéressante, 1400 gènes présentent une expression plus forte dans les tumeurs agressives que dans les tumeurs des patients de bon pronostic. Cette signature génique associée aux tumeurs de pronostic sombre se révèle particulièrement intéressante, étant donné qu'elle peut en théorie contenir des expressions géniques directement associées à l'agressivité des cellules cancéreuses, leur capacité à proliférer, à se disséminer ou à résister au traitement. Dans cette signature associée à l'agressivité tumorale, nous sommes partis à la recherche des expressions aberrantes de gènes tissu-spécifiques. Comme présenté plus tôt (**voir Introduction, paragraphe 3.3.4**), les expressions hors-contexte de gènes dont l'expression est normalement restreinte à un tissu particulier nous paraissaient être de bons candidats pour une étude plus poussée, pouvant permettre de proposer de nouveaux biomarqueurs de pronostic sombre, voire de nouvelles cibles thérapeutiques dans le traitement des tumeurs pulmonaires agressives. Nous nous sommes alors concentrés sur l'un des gènes surexprimés dans les tumeurs de patients de pronostic sombre, le gène *PRL* codant l'hormone hypophysaire prolactine, car son expression ectopique dans les tumeurs présente une association forte avec un pronostic sombre (**voir paragraphe 2.2.2**).

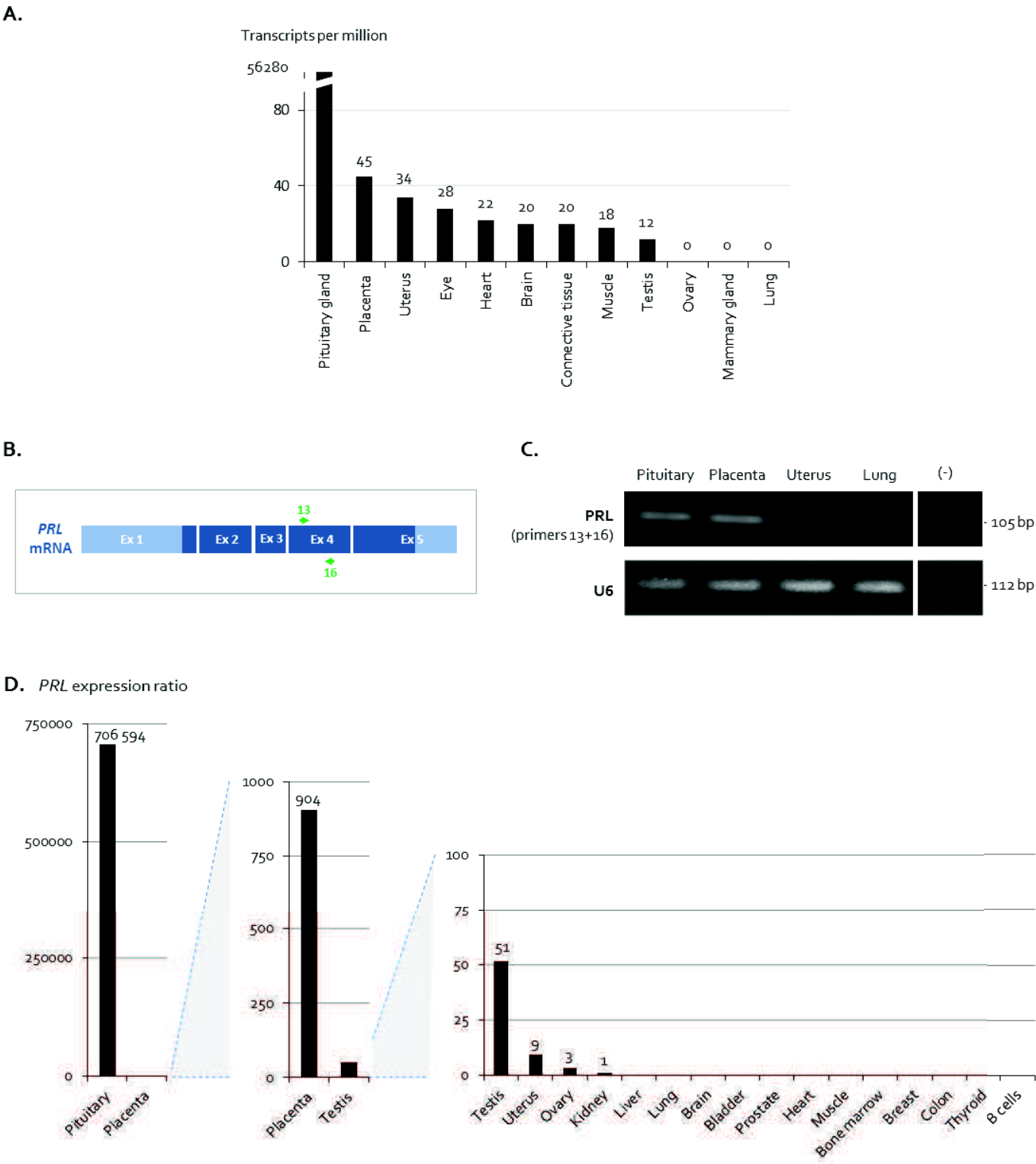


Figure 34: *PRL* gene expression in human tissues: predominant in pituitary gland, some extra-pituitary sources but not detected in non-tumoral lung

A. EST profile of *PRL* gene in human tissues (UniGene Hs. 1905). Approximate *PRL* gene expression pattern is inferred from EST counts in cDNA library sources.

Adapted from UniGene (NCBI): <http://www.ncbi.nlm.nih.gov/UniGene/ESTProfileViewer.cgi?uglist=Hs.1905>

B. Primers 13 and 16 used in PCR and qPCR experiments are specific of *PRL* gene and localized in exon 4. Various adult human tissues RNAs were submitted to RT. PCR and qPCR analysis were performed using RT products as template.

C. PCR analysis with specific primers in the *PRL* coding sequence on 4 adult human tissues cDNA.

U6 amplification is used as loading control. Negative control (-) is performed with H_2O .

D. Expression ratios of *PRL* gene in human tissues obtained from qPCR experiment.

PRL expression ratio is obtained by normalization with 2 different control genes (*Actin* and *U6*).

An expression ratio of 100 is assigned to the mean of *PRL* expression ratios in the six *PRL*-expressing tissues.

2.1.2. L'expression du gène *PRL*, restreinte à l'hypophyse et quelques sources extra-hypophysaires, n'est pas détectée dans le poumon sain

J'ai tout d'abord cherché à affiner le patron d'expression du gène *PRL* au sein des différents tissus humains (**Figure 34**). L'analyse des profils EST (Expressed Sequence Tag) disponibles en ligne pour le gène *PRL* (identifiant Unigene Hs.1905) confirme l'expression ultra-majoritaire du gène dans l'hypophyse humaine (**Figure 34A**). D'autres tissus apparaissent comme sites secondaires d'expression du gène *PRL*, dont le placenta, l'utérus ou le testicule. On peut noter l'absence d'EST *PRL* dans le poumon sain. Ce profil d'expression du gène *PRL* a été confirmé expérimentalement par des expériences de PCR et de qPCR sur différents tissus humains sains. Après reverse-transcription des ARN totaux de tissus, les cDNA ont été soumis à une amplification grâce à une paire de primers spécifiques de l'exon 4 du gène *PRL* (**Figure 34B**). En PCR, on détecte une amplification dans l'hypophyse et le placenta, mais ni dans l'utérus, ni dans le poumon (**Figure 34C**). Une quantification des niveaux d'expression est permise par une PCR quantitative (**Figure 34D**). L'expression est ultra-majoritaire dans l'hypophyse. Parmi les sources extra-hypophysaires, c'est le placenta qui exprime le plus fortement le gène *PRL*, suivi du testicule. De faibles expressions sont détectées dans l'utérus, l'ovaire et le rein, alors qu'aucune amplification n'est obtenue dans le poumon non pathologique. Ces résultats confirment que le gène *PRL*, sans être un gène tissu-spécifique *sensu stricto*, a bien une expression restreinte à l'hypophyse et à quelques sources extra-hypophysaires n'incluant pas le poumon adulte sain.

2.1.3. Le gène *PRL* est exprimé de manière aberrante dans 10% des tumeurs pulmonaires

Etant donné que le gène *PRL* n'est pas exprimé dans le poumon sain mais qu'il a été identifié dans la signature génique des tumeurs pulmonaires agressives, nous avons étudié l'expression du gène *PRL* dans les cancers du poumon (**Figure 35**). Nous nous sommes tout d'abord penchés sur les données transcriptomiques des patients de la cohorte (**Figure 35A et B**). Le niveau d'expression du gène *PRL*, dans les tumeurs de patients mais également dans 112 tissus somatiques sains analysés conjointement, découle du signal d'un ensemble de sondes (probeset 205445_at) spécifiques de la séquence nucléotidique du gène *PRL*, toutes situées dans les exons 4 et 5 des transcrits *PRL* (**Figure 35A**). La spécificité de ce probeset est validée par l'analyse sur 112 tissus somatiques, étant donné que le signal le plus fort est détecté dans les 3 échantillons d'hypophyse (données non présentées). Pour analyser les données transcriptomiques des patients, nous avons été confrontés à la difficulté d'établir un seuil (**Figure 35B**). En effet, les données transcriptomiques, qui sont issues d'un signal d'hybridation sur puce soumis au bruit de fond, ne peuvent *a priori* pas être interprétées en terme d'expression ou d'absence d'expression du gène. Nous avons alors établi un seuil en-dessous duquel on considère que le gène *PRL* n'est pas exprimé, en moyennant la valeur du signal du probeset *PRL* dans 109 échantillons de tissus humains (les 3 échantillons hypophysaires ayant été exclus). On observe alors que parmi les 293 patients de la cohorte, 27 patients exhibent au sein de leur tumeur une expression du gène *PRL* qui dépasse ce seuil. On considère que le gène *PRL* est exprimé de manière aberrante dans ces 27 tumeurs et les patients correspondants (9,2% de la cohorte) sont renommés « patients *PRL*-positifs » (**Figure 35B**). Le seuil que nous avons choisi d'établir nous paraît pertinent au vu de la distribution des ratios d'expression *PRL* dans les 293 tumeurs : lorsque les 293 patients sont classés par ordre croissant d'expression, le niveau d'expression du gène *PRL* est quasi-constant dans les tumeurs considérées comme *PRL*-négatives, alors que les 27 patients *PRL*-positifs se situent après un point d'inflexion (**Figure 35B, panneau de droite**). J'ai alors souhaité confirmer partiellement ces résultats en PCR sur les cDNA de quelques patients de la cohorte (**Figure 35C**). En utilisant des primers spécifiques du gène *PRL*, on obtient une amplification pour 3 échantillons (patients n°56, 59 et 62). Ces trois patients ont bel et bien été considérés comme *PRL*-positifs dans l'analyse des données transcriptomiques, au contraire des 3 autres patients considérés comme *PRL*-négatifs. Ces résultats dans leur ensemble mettent en évidence que le gène *PRL*, normalement réprimé dans le poumon non-pathologique, est exprimé de manière aberrante dans environ 10% des tumeurs pulmonaires.

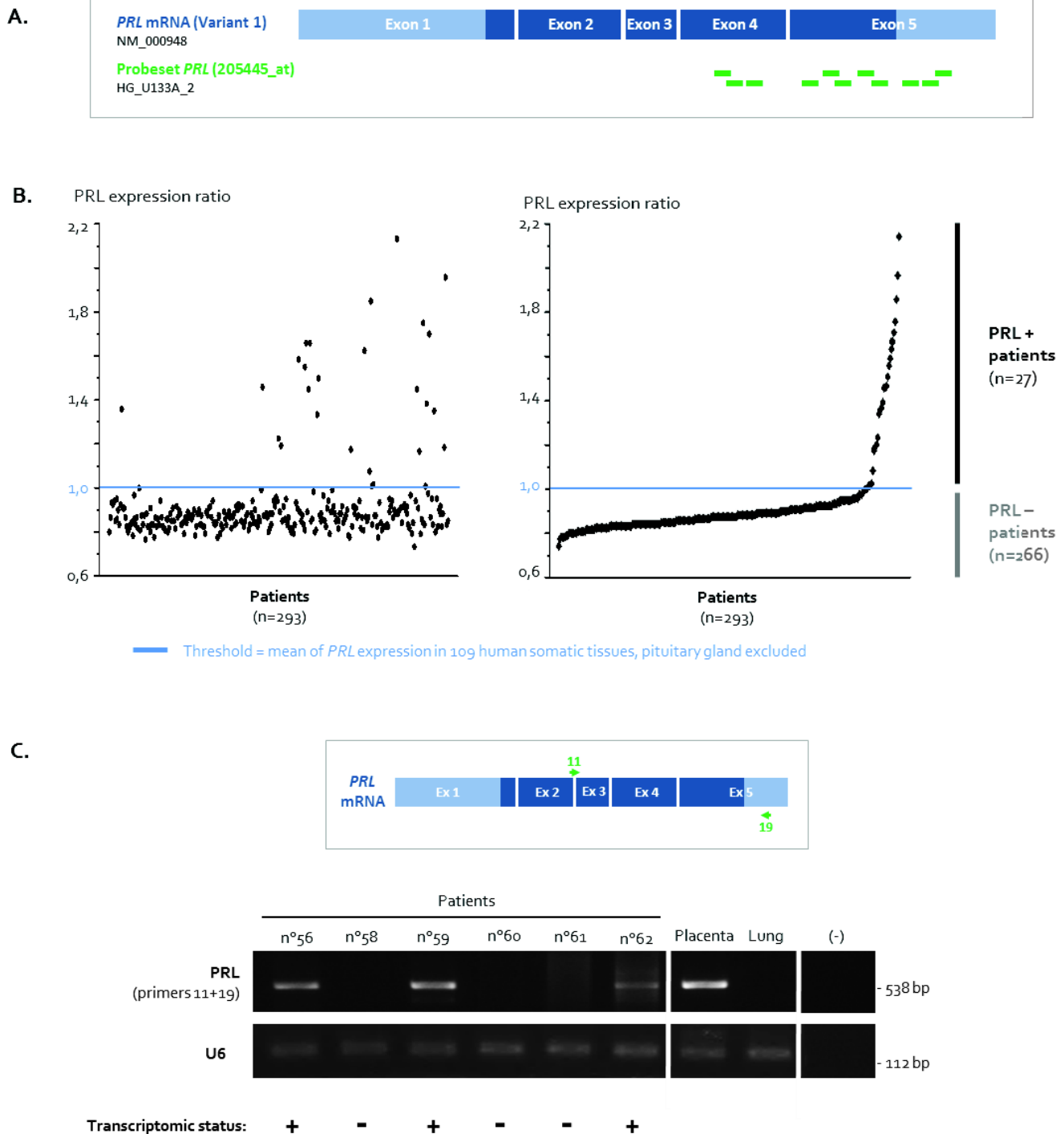


Figure 35: *PRL* gene is expressed within the pulmonary tumor of 9.2% patients

A. The 205445_at probeset is composed of 11 probes recognizing exons 4 and 5 of *PRL* gene sequence.

Adapted from Germ Online ContigView (with OLIGO_HG_U133_2 features):

http://www.germonline.org/Homo_sapiens/geneview?altsplice=%7Chg_u133a_2%3Aon&db=core&gene=ENSG00000172179

B. Expression ratios of *PRL* probe set (205445_at) obtained by the transcriptomic analysis (Affymetrix technology) of 293 lung tumors of the cohort. Expression ratios were obtained by normalizing RMA signal data exported from Gene Spring with the mean of *PRL* signal in 109 human somatic tissues.

We consider that *PRL* gene is expressed within the tumor if the *PRL* expression ratio is higher than the threshold (blue line). The threshold is established by averaging the *PRL* signal in 109 *PRL*-non expressing human tissues (pituitary gland excluded).

27 tumors (9.2% of the cohort) exhibiting a *PRL* expression ratio higher than the threshold are considered as *PRL*-expressing tumors: patients are renamed *PRL*-positive patients (*PRL* + patients).

(Left panel) Patients are randomly assigned. (Right panel) Patients are classified in ascending order of *PRL* expression ratios.

C. PCR analysis with *PRL* specific primers on 6 patients cDNA. *U6* amplification is used as loading control. Negative control (-) is performed with H₂O. The *PRL* status (+: expressing or -: not expressing) as determined by transcriptomic analysis is indicated for each patient.

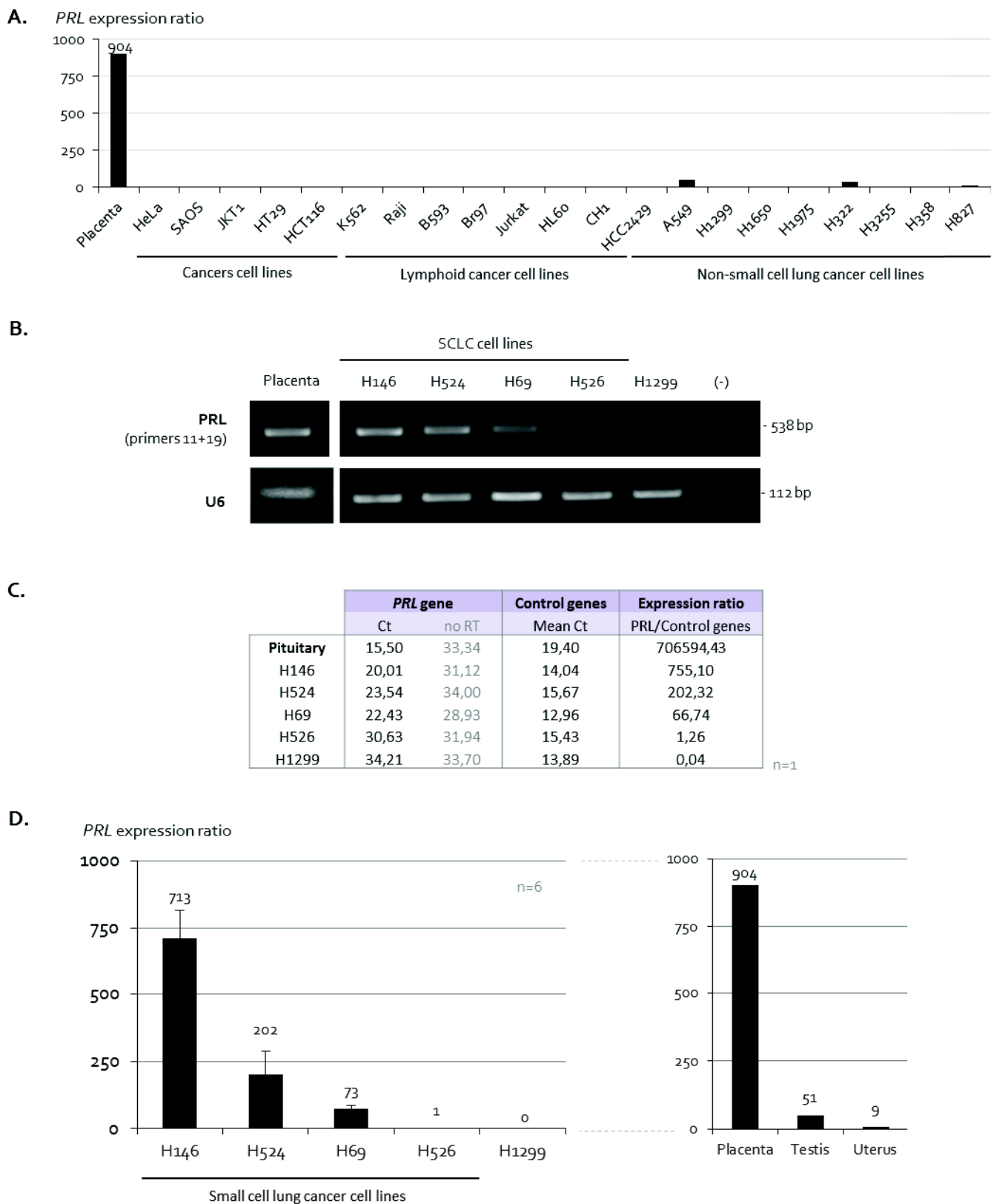


Figure 36: Expression pattern of *PRL* gene in human cancer cell lines: identification of *PRL* expressing-small cell lung cancer cell lines

Various human cancer cells lines RNAs were submitted to RT.

PCR and qPCR analysis with specific *PRL* primers were performed using RT products as template.

A. qPCR analysis on various human cancer cells cDNA.

B. PCR analysis with specific primers in the *PRL* coding sequence on 5 human lung cancer cells cDNA.

C. qPCR analysis on 5 human lung cancer cell lines. Example of raw data obtained.

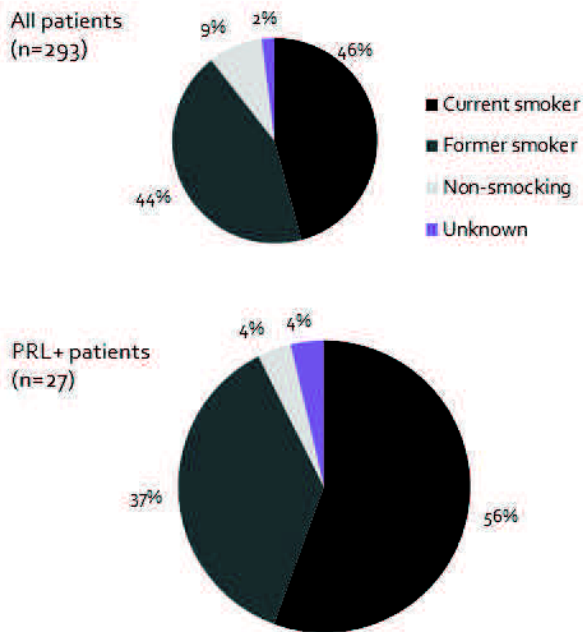
D. Expression ratios of *PRL* gene in 5 human lung cancer cell lines (and comparison with tissues *PRL* expression ratios) obtained from 6 independent RT-qPCR experiments, with normalization thanks to 2 different control genes (*Actin* and *U6*).

100% = Mean of *PRL* expression ratios from the six *PRL*-expressing tissues. Error bars represent the standard deviation.

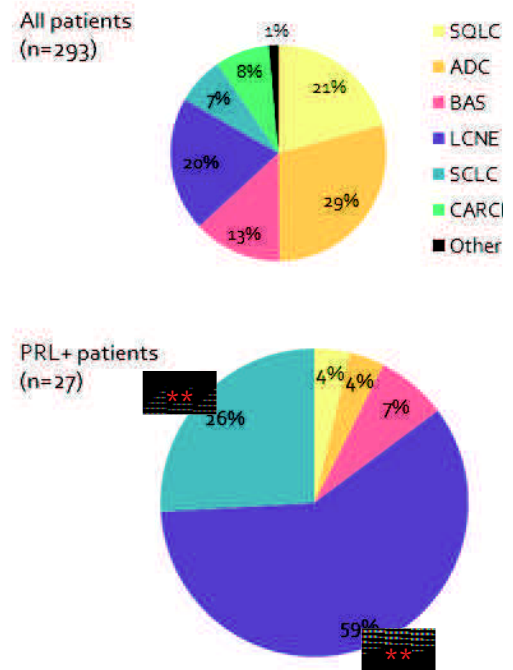
A.

	PRL + patients	PRL - patients	All patients
Number of cases	27	266	293
Men	96,3%	83,8%	85,0%
Mean age at diagnosis	59,33	61,11	60,95
Mean age at surgery	59,78	61,69	61,51
Mean survival (months)	30,22	70,62	66,9

B. Smoking status



C. Histological subtype of the tumor



D. TNM classification

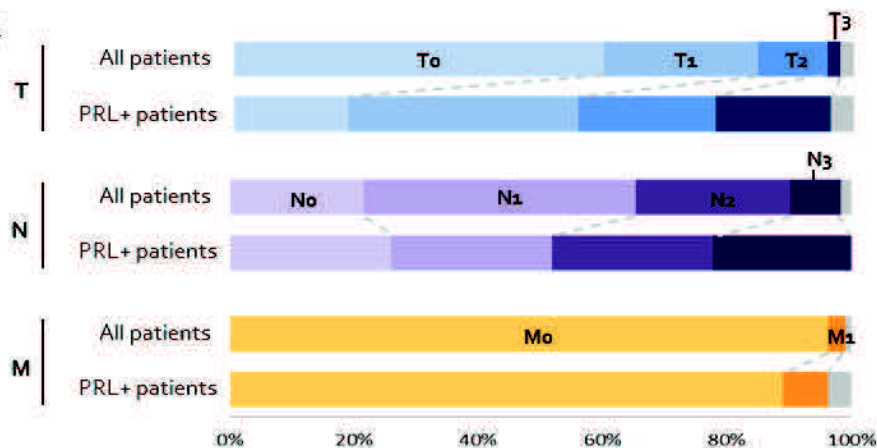


Figure 37: Characterization of PRL + patients

A. Comparison of general data between all patients (n=293) and PRL+ patients (n=27).

B. Comparison of the smoking status between all patients (n=293) and PRL+ patients (n=27).

C. Comparison of the histological type of the tumor between all patients (n=293) and PRL+ patients (n=27).

Distribution enrichment was tested using the hypergeometric test : ** p value < 0.01.

SQLC: squamous lung carcinoma ; ADK: adenocarcinoma ; BAS: basaloid carcinoma ; LCNE: large cell neuroendocrine lung carcinoma ; SCLC : small cell lung carcinoma ; CARCI : carcinoid lung tumor

D. Comparison of TNM values distribution in all patients (n=293) and PRL+ patients (n=27).

T: primary tumor size ; N: spread to regional lymph nodes ; M: presence of distant metastasis

2.1.4. Le gène *PRL* est exprimé de manière aberrante par des lignées de cancer du poumon à petites cellules

Pour confirmer l'expression ectopique du gène *PRL* dans les cellules cancéreuses pulmonaires et obtenir un modèle cellulaire sur lequel travailler au laboratoire, je suis partie à la recherche d'une lignée cellulaire cancéreuse humaine qui exprimerait le gène *PRL* (**Figure 36**). Nos premières analyses par qPCR sur des lignées de cancers variés (cancer du poumon, mais aussi cancers immuns ou du côlon) se sont révélées infructueuses (**Figure 36A**). Puis nous nous sommes procuré 4 nouvelles lignées de cancer de poumon, toutes issues de carcinomes pulmonaires à petites cellules (CBPC) : H146, H524, H69 et H526. Les trois premières présentent une expression du gène *PRL* détectée en PCR (**Figure 36B**). L'analyse en PCR quantitative nous permet d'observer que le niveau d'expression de *PRL* dans ces cellules est faible, bien en deçà du niveau hypophysaire, mais toutefois relativement proche du niveau observé dans le placenta (**Figure 36D**). Les lignées H146, H524 et H69 expriment donc de manière aberrante le gène *PRL*. La lignée H526 quant à elle présente une très légère amplification en qPCR qui semble non spécifique (valeur de Ct proche du contrôle No RT) (**Figure 36C**). Elle sera donc considérée *PRL*-négative, tout comme la lignée de cancer du poumon non-à petites cellules H1299.

2.2. L'expression du gène *PRL* dans les cellules cancéreuses pulmonaires est associée à un pronostic sombre et à une chimiorésistance accrue

2.2.1. Les tumeurs qui expriment le gène *PRL* sont principalement des tumeurs neuroendocrines à caractéristiques de tumeurs agressives

Dans le but de décrypter les conséquences de l'expression hors-contexte du gène *PRL* dans les cellules cancéreuses pulmonaires, j'ai tout d'abord cherché à préciser les caractéristiques des tumeurs *PRL*⁺ en analysant les données cliniques des patients, le type histologique et le stade de la tumeur (**Figure 37**). Les patients *PRL*⁺ présentent un sex-ratio encore plus déséquilibré que les patients *PRL*⁻ : il n'y a qu'une seule femme au sein des 27 patients *PRL*⁺ identifiés (**Figure 37A**). Le diagnostic de cancer du poumon est réalisé un peu plus précocement, et la survie moyenne des patients *PRL*⁺ est diminuée de plus de moitié par rapport aux patients *PRL*⁻. On note également une très légère sur-représentation des fumeurs au sein du groupe de patients *PRL*⁺ (**Figure 37B**). De manière intéressante, la répartition des sous-types histologiques des tumeurs *PRL*⁺ est très différente de celle de l'ensemble des tumeurs. On note un biais significatif de la distribution (Loi hypergéométrique : p value < 0,01) vers les tumeurs à caractéristiques neuroendocrines, qui représentent au total 85% des tumeurs *PRL*⁺, qu'il s'agisse de carcinomes bronchiques à petites cellules SCLC (26% contre 7%) ou de carcinomes neuroendocrines à grandes cellules LCNE (59% contre 20%) (**Figure 37C**). Enfin, la comparaison des paramètres T (taille de la tumeur), N (envahissement ganglionnaire) et M (présence de métastases) semble indiquer que les tumeurs *PRL*⁺ sont des tumeurs présentant une croissance et une invasivité plus marquée que les tumeurs *PRL*⁻. Se pose alors la question de savoir si, comme le laissent penser ces données, les tumeurs *PRL*⁺ sont effectivement plus agressives.

2.2.2. L'expression aberrante du gène *PRL* au sein de la tumeur est associée à un pronostic sombre pour le patient

Nous avons réalisé des analyses de survie, basées sur les données cliniques recueillies sur les patients de la cohorte, permettant de comparer les pronostics des patients *PRL*⁺ et *PRL*⁻. Les courbes de survie, réalisées selon la méthode de Kaplan-Meier ([Kaplan, 1958](#)), représentent la probabilité de survie des patients (en ordonnée) au cours du temps de suivi de la cohorte (60 mois, en abscisse). Les deux sous-groupes de patients *PRL*⁺ et *PRL*⁻ présentent des pronostics très différents (**Figure 38A**). La survie est diminuée de manière significative (p -value < 0,0001) chez les patients dont la tumeur exprime le gène *PRL*. Après 12 mois, moins de

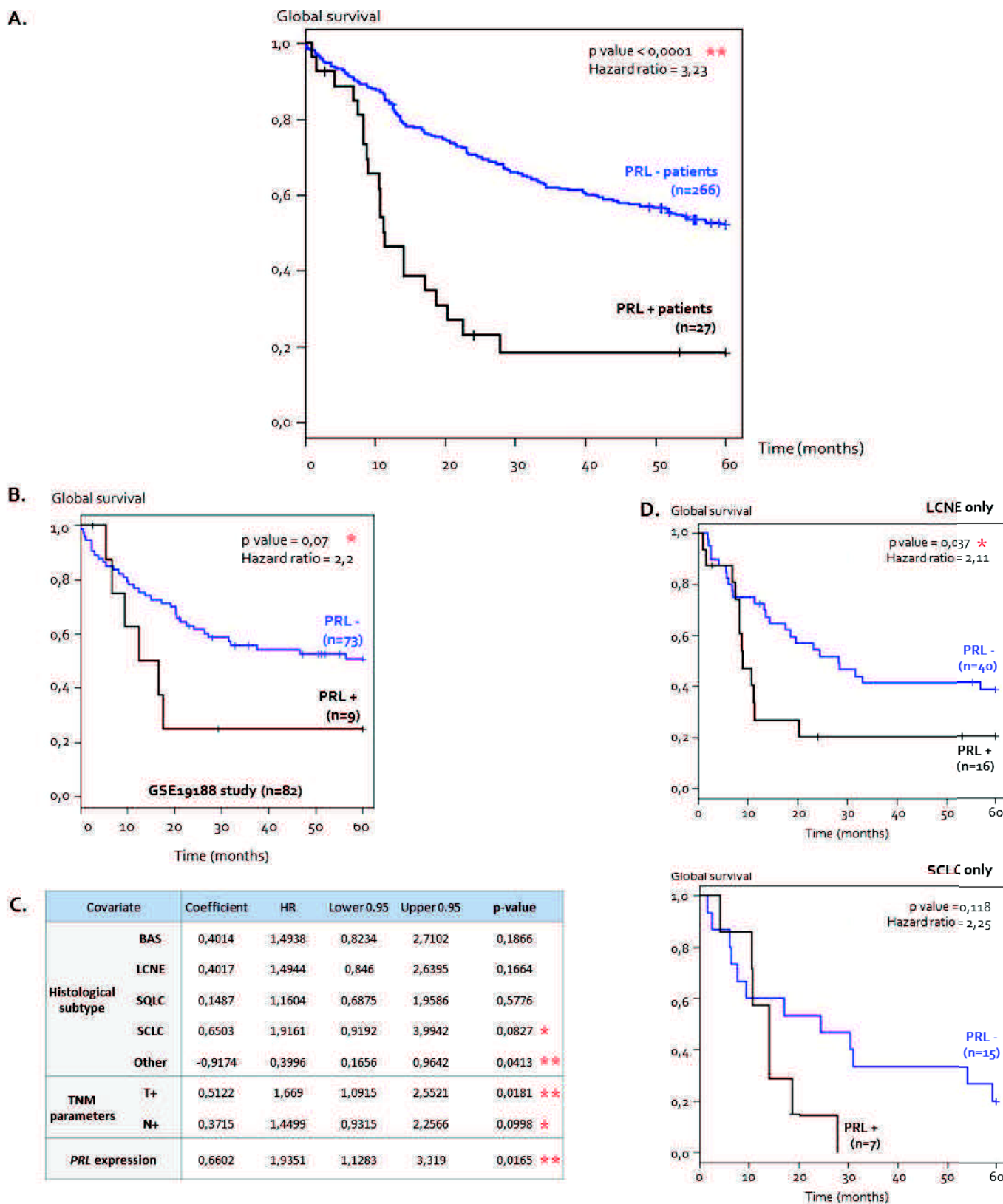


Figure 38: *PRL* gene aberrant expression within a pulmonary tumor is associated with a patient's poor prognosis

A. Survival probability curves (estimated by the Kaplan-Meier method) for *PRL*- (n=266) and *PRL*+ patients (n=27). The significance of the difference in prognosis is assessed by the univariate logrank test. The hazard ratio estimates the difference in death rates between the 2 groups.

B. Validation performed on the GSE19188 external study (n=82 patients with lung cancer).

C. Multivariate analysis according to the Cox model. Following covariates are incorporated: histological subtype (BAS, LCLC, SQLC, SCLC, other), TNM classification (T and N parameters) and *PRL* expression.

D. Survival probability curves (estimated by the Kaplan-Meier method) for *PRL*- and *PRL*+ patients, with large cell lung cancer only (n=56) (top) or small cell lung cancer only (n=22) (bottom).

** p value < 5%

* p value < 10%

50% des patients *PRL*⁺ sont toujours en vie, contre plus de 80% chez les patients *PRL*⁻. Plus généralement, les patients *PRL*⁺ présentent 3 fois plus de risque de mourir de leur cancer du poumon que les patients *PRL*⁻ (Hazard ratio = 3,23). Ce résultat a été validé (avec une p-value inférieure à 10%) dans une étude externe réunissant 82 patients souffrant de cancers pulmonaires, dont 9 individus *PRL*⁺ (**Figure 38B**).

La question se pose alors de savoir si cette survie réduite des patients *PRL*⁺ est directement liée à l'expression du gène *PRL*, ou si elle n'est que le reflet de l'enrichissement du sous-groupe *PRL*⁺ en carcinomes bronchiques à petites cellules, connus pour être les formes les plus agressives de tumeurs pulmonaires (**voir Introduction, paragraphe 2.2.2.1.2**). La réponse est donnée par une analyse multivariée (**Figure 38C**). Celle-ci permet d'étudier de manière indépendante le poids de plusieurs paramètres (type histologique, paramètres TNM ou expression du gène *PRL*) dans la réduction de survie observée. Cette analyse montre que c'est bien l'expression du gène *PRL* qui constitue le paramètre déterminant dans le pronostic sombre des patients. En effet, parmi les différents paramètres testés, c'est l'expression de *PRL* qui est associée à la p-value la plus faible. On peut illustrer ce résultat en étudiant la probabilité de survie des patients présentant le même sous-type histologique de tumeur (**Figure 38D**). Lorsque l'on ne considère que les tumeurs neuroendocrines à grandes cellules de la cohorte (56 patients) (**panneau du haut**), on retrouve une survie significativement diminuée chez les patients *PRL*⁺. La même tendance est observée lorsque l'on considère uniquement les tumeurs à petites cellules (**panneau du bas**), même si la significativité statistique n'est pas atteinte du fait de la faible taille de l'échantillon. Ainsi, au sein de ces tumeurs à petites cellules très agressives, l'expression aberrante du gène *PRL* est encore capable de discriminer un sous-groupe de patients *PRL*⁺ au pronostic particulièrement sombre.

2.2.3. L'expression aberrante du gène *PRL* confère aux cellules cancéreuses pulmonaires une plus grande résistance à un stress génotoxique

J'ai montré que les tumeurs *PRL*⁺ sont particulièrement agressives. Je me suis alors demandé si l'on pouvait démontrer un rôle direct de l'expression du gène *PRL* dans l'agressivité des cellules CBPC *PRL*-positives identifiées précédemment (H146, H524 et H69). Pour tester l'implication de l'expression de *PRL* dans le phénotype cancéreux, j'ai tout d'abord mis au point une extinction du gène *PRL* dans les lignées CBPC *PRL*⁺ grâce à la technique d'ARN interférence (**Figure 39A**). La transfection de deux couples de siARN permet de réduire de manière significative le niveau d'expression du gène *PRL* dans les 3 lignées *PRL*⁺. Le silençage le plus efficace est observé avec le duplex siARN n°1, qui permet de réduire l'expression de *PRL* de plus de 90% dans la lignée H146. J'ai alors cherché à savoir quel pouvait être l'impact de ce silençage du gène *PRL* sur les paramètres cellulaires pouvant jouer un rôle dans l'agressivité des cellules tumorales, en particulier la prolifération, la mort cellulaire et l'apoptose induite par une drogue (**Figures 39 et 40**).

Dans un premier temps, je n'ai noté aucun effet de la baisse d'expression de *PRL* sur la morphologie des cellules CBPC *PRL*⁺. Aucun changement de la morphologie ou de l'adhérence des cellules n'est observable en microscopie après transfection des siARN *PRL* sur la lignée H146 qui pousse sous forme d'amas en suspension (**Figure 39B**). J'ai également cherché à déterminer un effet du siARN *PRL* sur la prolifération cellulaire. La croissance en amas rendant impossible des comptages cellulaires précis pour réaliser une cinétique de croissance, j'ai étudié le cycle cellulaire des lignées CBPC en cytométrie de flux (**Figure 39C**). Malgré une légère tendance à la baisse du nombre de cellules H146 en phase S sous siARN *PRL*, aucun effet significatif du silençage *PRL* n'a pu être mis en évidence sur le cycle dans les différentes lignées.

Je me suis alors intéressée à la mort cellulaire des lignées CBPC soumises au silençage du gène *PRL*, en mesurant à la fois à la mortalité basale en condition contrôle et à la mortalité induite par un stress (**Figure 40**). Les mortalités basale et induite par le traitement par la drogue Actinomycine D ont été quantifiées par cytométrie de flux après marquage des cellules apoptotiques. On note tout d'abord que la baisse d'expression de *PRL* n'a pas d'effet sur la mortalité basale des cellules en absence de traitement (**en gris, Figure 40A, panel du milieu**). La mortalité des cellules H146 et H524 reste ainsi stable aux alentours de 10% dans un milieu de

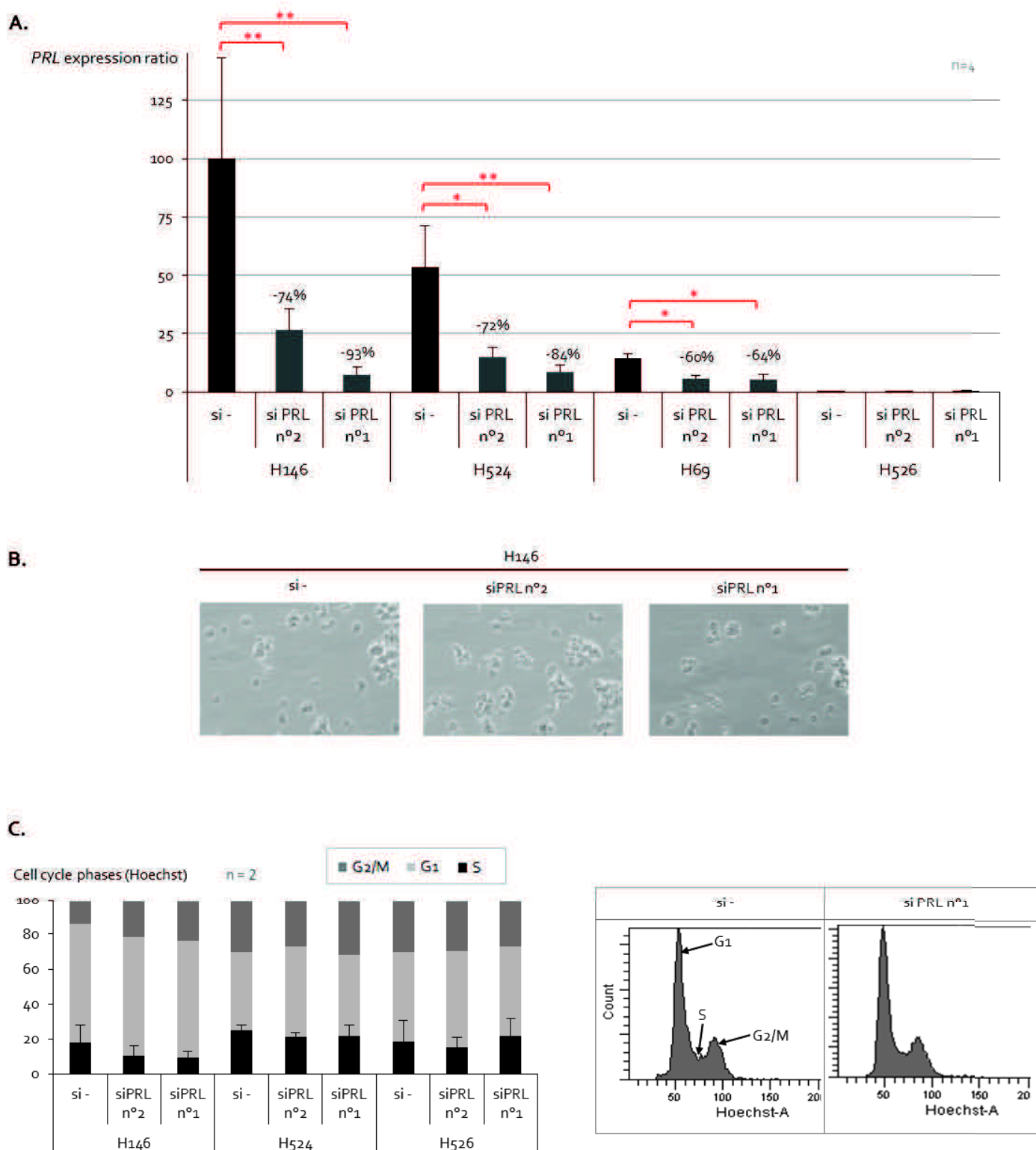


Figure 39: Effects of *PRL* silencing by RNA interference on *PRL*+ SCLC cell lines

A. SCLC cell lines were transfected with siRNA scrambled (si -) or siRNA targeting *PRL* mRNAs (siPRL, 2 different siRNA duplex) for 48 hours. Expression ratios of *PRL* gene obtained from 4 independent RT-qPCR experiments. Expressions are normalized thanks to *Actin* control gene. 100% = *PRL* expression ratio in H146 si -. Silencing efficiencies are indicated. The significance of the difference in *PRL* expression between si - condition and siPRL conditions is assessed by a Student t-test of mean comparison

* p value < 10% ** p value < 5%

B. Effect of *PRL* silencing on cell morphology of the H146 cell line.

C. Effect of *PRL* silencing on cell cycle. The cell cycle phase of each cell has been assessed by flow cytometry with an Hoechst staining of DNA. (Left panel) Quantification of cell cycle phases on 2 independent experiments. The significance of the difference in cell cycle phase between si - condition and siPRL conditions is assessed by a Student t-test of mean comparison.

(Right panel) Example of flow cytometry cell cycle analysis on the H524 cell line. On the y-axis is the number of cells, on the x-axis the fluorescence of Hoechst. Corresponding cell cycle phase are indicated by arrows.

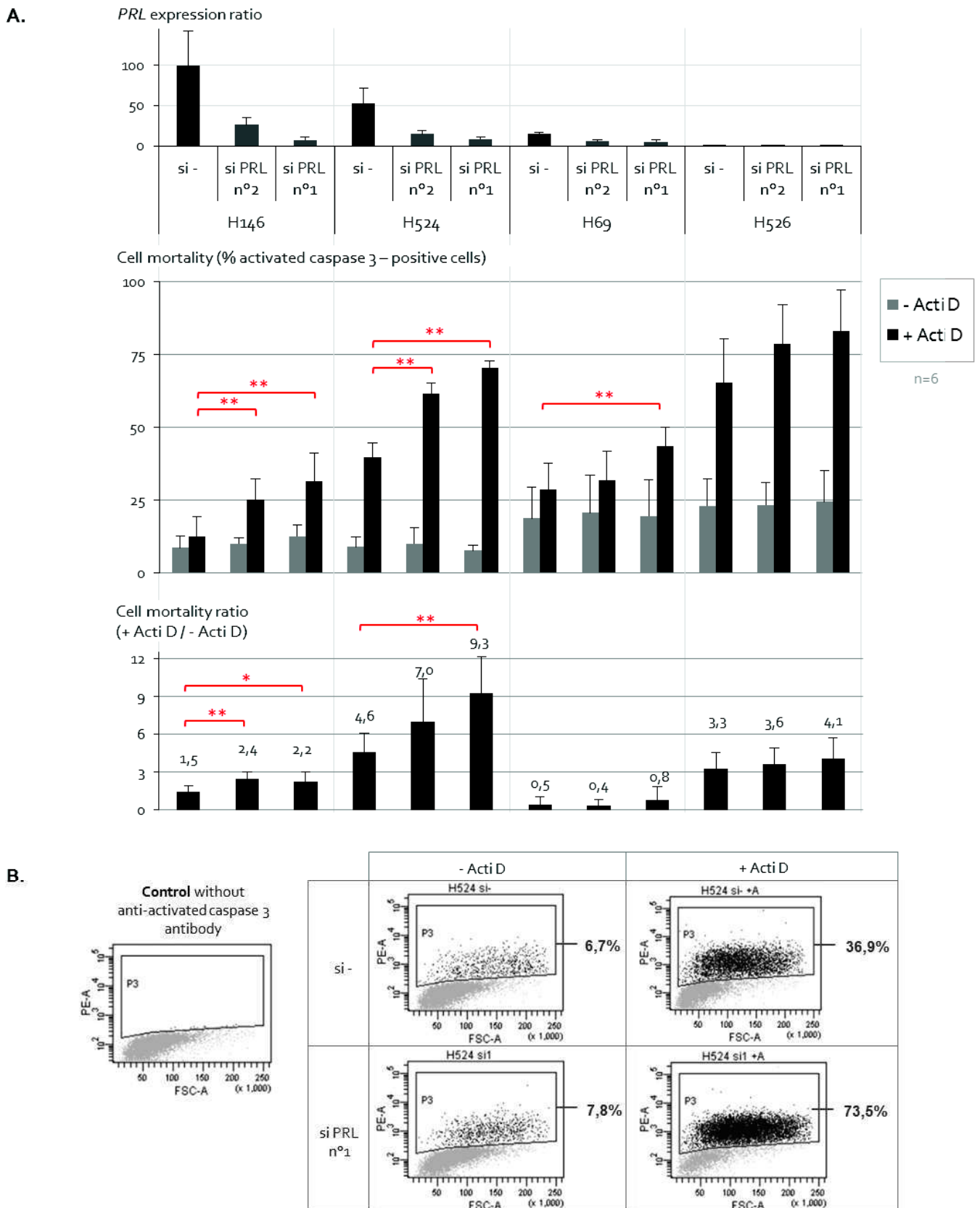


Figure 40: Effect of *PRL* silencing on cell mortality in *PRL*+ SCLC cell lines

SCLC cell lines were transfected with siRNA scrambled (si -) or targeting *PRL* (siPRL, 2 different siRNA duplex) for 48 hours, and treated for 12 hours by 250 mM Actinomycin D (Acti D) or control treatment.

A. (Top panel) Silencing of *PRL* gene as assessed by RT-qPCR. (Middle panel) Cell mortality was measured by flow cytometry after incubation with an anti-activated Caspase 3 antibody in 6 independent experiments. (Bottom panel) Cell mortality ratio (Acti D-induced cell death versus basal cell death in control condition) was calculated for each of the 6 experiments.

The significance of the difference in cell mortality between si - condition and siPRL conditions is assessed by a Student t-test of mean comparison. * p value < 10% ** p value < 5%

B. Example of flow cytometry data. Each dot represents one H524 cell. On the y-axis is represented the PE fluorescence of the anti-activated caspase 3 staining. On the x-axis is the FSC size parameter. P3 region comprises apoptotic cells.

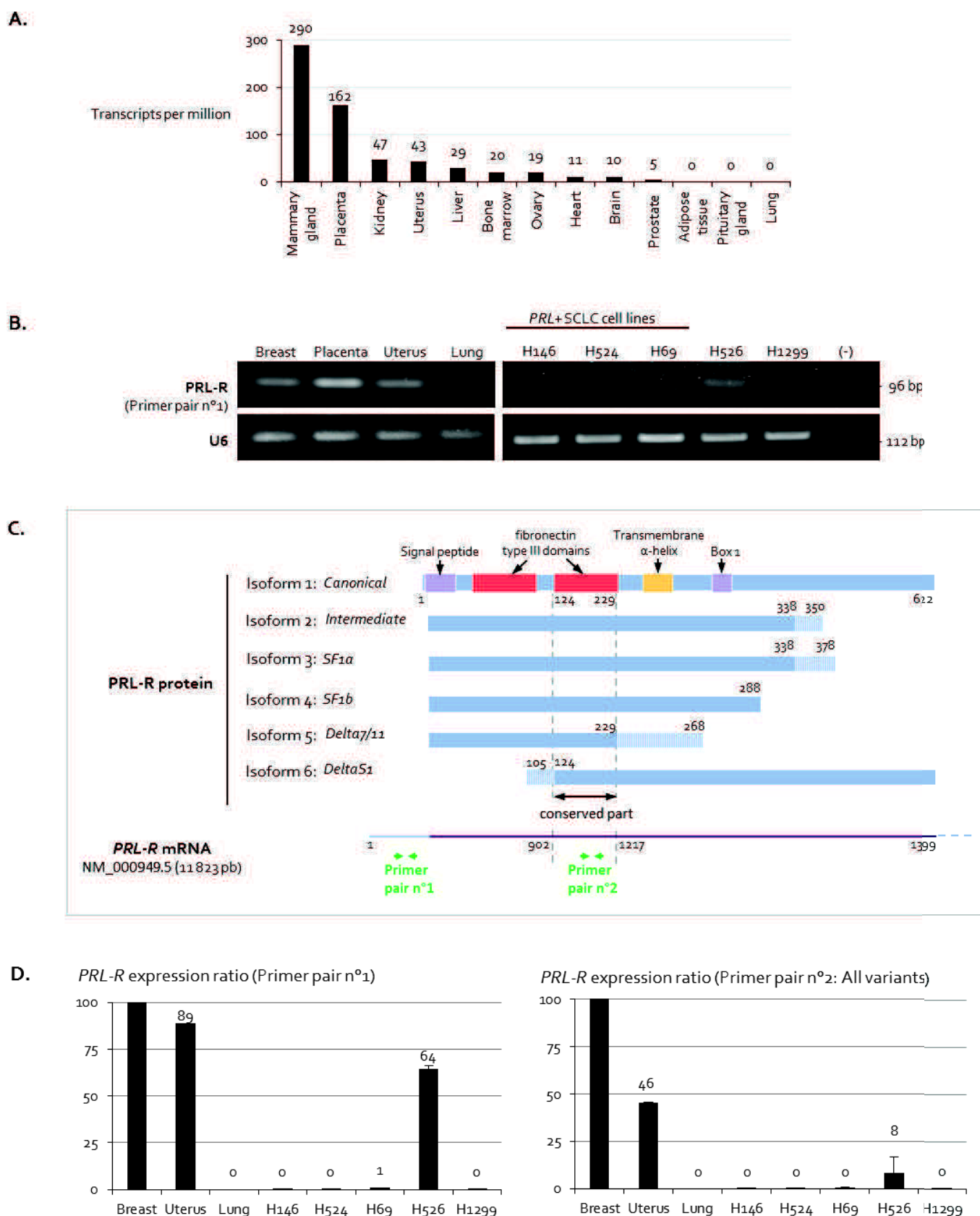


Figure 41: Expression pattern of *PRL-R* gene in human tissues and in lung cancer cell lines

A. EST profile of *PRL-R* gene in human tissues (UniGene Hs. 368587).

Adapted from UniGene (NCBI): <http://www.ncbi.nlm.nih.gov/UniGene/ESTProfileViewer.cgi?uglist=Hs.368587>

B. PCR and qPCR analysis with specific *PRL-R* primers were performed on 4 human tissues and 5 human lung cancer cells lines. U6 amplification is used as loading control. Negative control (-) is performed with H₂O.

C. Primers recognizing all variants of *PRL-R* mRNA have been designed in the sequence coding for a common protein domain.

D. Expression ratios of *PRL-R* gene in human tissues and lung cell lines were assessed by qPCR with primers that only recognize the variant 1 of *PRL-R* mRNA coding for the canonical PRL-R protein (left panel) and primers that recognize all *PRL-R* mRNA variants (right panel). Expressions are normalized thanks to the *Actin* control gene. 100% = *PRL-R* expression ratio in the breast.

culture sans Actinomycine D, et ce quel que soit le niveau d'expression de *PRL* (conditions si -, siPRL n°1 ou n°2). L'application d'un stress génotoxique, sous la forme de l'ajout d'actinomycine D dans le milieu de culture des cellules, induit une augmentation de la mortalité (**en noir, Figure 40A, panel du milieu**). Pour la lignée H524 par exemple, la mortalité Actinomycine-induite atteint 40% pour les cellules non transfectées. Suite à l'application de cette drogue, on constate une différence significative et reproductible de la mortalité entre les cellules non-transfectées (condition si -) et les cellules chez qui l'expression du gène *PRL* est réduite par siARN (conditions siPRL n°1 et n°2). Par exemple, lorsque l'expression du gène *PRL* est silencée chez les cellules H524, la mortalité sous Actinomycine D dépasse les 70%, contre 40% lorsque le gène *PRL* s'exprime normalement (**Figure 40B**). Ainsi, alors qu'un traitement par la drogue multiplie en moyenne la mortalité basale des cellules H524 par quatre (condition si -), la suppression concomitante de l'expression de *PRL* majore l'effet de la drogue (multiplication de la mortalité par 9 après transfection du siPRL n°1) (**Figure 40A, panneau du bas**). On peut noter que plus le silençage du gène *PRL* est efficace (siARN n°1 plus efficace que le siARN n°2), plus l'augmentation de la mortalité est marquée. On remarque également qu'aucune augmentation significative de la mortalité Actinomycine D-induite n'est observée suite à la transfection des siARN dans la lignée H526 *PRL*-négative. Ces résultats montrent que lorsque l'expression du gène *PRL* est altérée, les cellules sont plus sensibles à un stress génotoxique, et démontrent que l'expression de *PRL* est directement impliquée dans la chimiorésistance des cellules CBPC *PRL*+ en culture.

Ces résultats dans leur ensemble démontrent une association directe entre l'expression aberrante du gène hypophysaire *PRL* par les cellules cancéreuses et la résistance de ces cellules à un stress. Ceci est à mettre en relation avec la plus grande agressivité des tumeurs pulmonaires observée chez les patients *PRL*+, agressivité qui pourrait être directement due à l'acquisition par les cellules tumorales d'une plus grande résistance aux différents stress subis (hypoxie, surveillance immune, traitements...). Se pose alors la question de savoir comment est médié cet effet de l'expression de *PRL* dans les cellules cancéreuses pulmonaires.

2.3. L'action de la prolactine est indépendante de son récepteur

Des données de la littérature soutiennent l'existence, au sein de tumeurs du sein mais aussi de la prostate, d'une sécrétion locale de prolactine, agissant via son récepteur dans une boucle autocrine/paracrine pour promouvoir la croissance tumorale (Clevenger et al., 1995 ; Dagvadorj et al., 2007) (Voir Discussion, paragraphe 3.2). Au vu de ces données, nous avons dans un premier temps retenu l'hypothèse suivante : l'expression du gène *PRL* dans les tumeurs bronchiques conduirait à une production locale de prolactine sécrétée par les cellules cancéreuses pulmonaires. Cette prolactine pourrait alors agir au sein de la tumeur, via son récepteur, dans une boucle autocrine/paracrine qui stimulerait la résistance des cellules.

2.3.1. Le récepteur à la prolactine n'est exprimé ni dans le poumon non-tumoral, ni dans les cellules cancéreuses qui expriment le gène *PRL*

Je me suis tout d'abord intéressée au patron d'expression du gène *PRL-R* codant le récepteur à la prolactine dans divers tissus humains sains (**Figure 41A**). L'analyse des profils EST disponibles en ligne pour le gène *PRL-R* (identifiant Unigene Hs.368587) confirme l'expression majoritaire du gène dans la glande mammaire. D'autres tissus humains expriment le gène *PRL-R* à des niveaux plus faibles, dont le placenta, l'utérus ou la prostate. On peut noter l'absence d'EST *PRL-R* dans le poumon sain. Ce résultat est confirmé par une analyse en PCR (**Figure 41B**). En utilisant des primers spécifiques de la séquence *PRL-R*, on obtient une amplification dans le sein, le placenta et l'utérus, mais pas dans le poumon non-pathologique. Se pose alors la question de savoir si le récepteur à la prolactine pourrait être, tout comme son ligand, exprimé de manière aberrante dans les cellules cancéreuses pulmonaires qui expriment le gène *PRL*. Les PCR sur 5 lignées cellulaires de poumon (**Figure 41B**) ne m'ont permis d'observer une expression de *PRL-R* que dans l'une

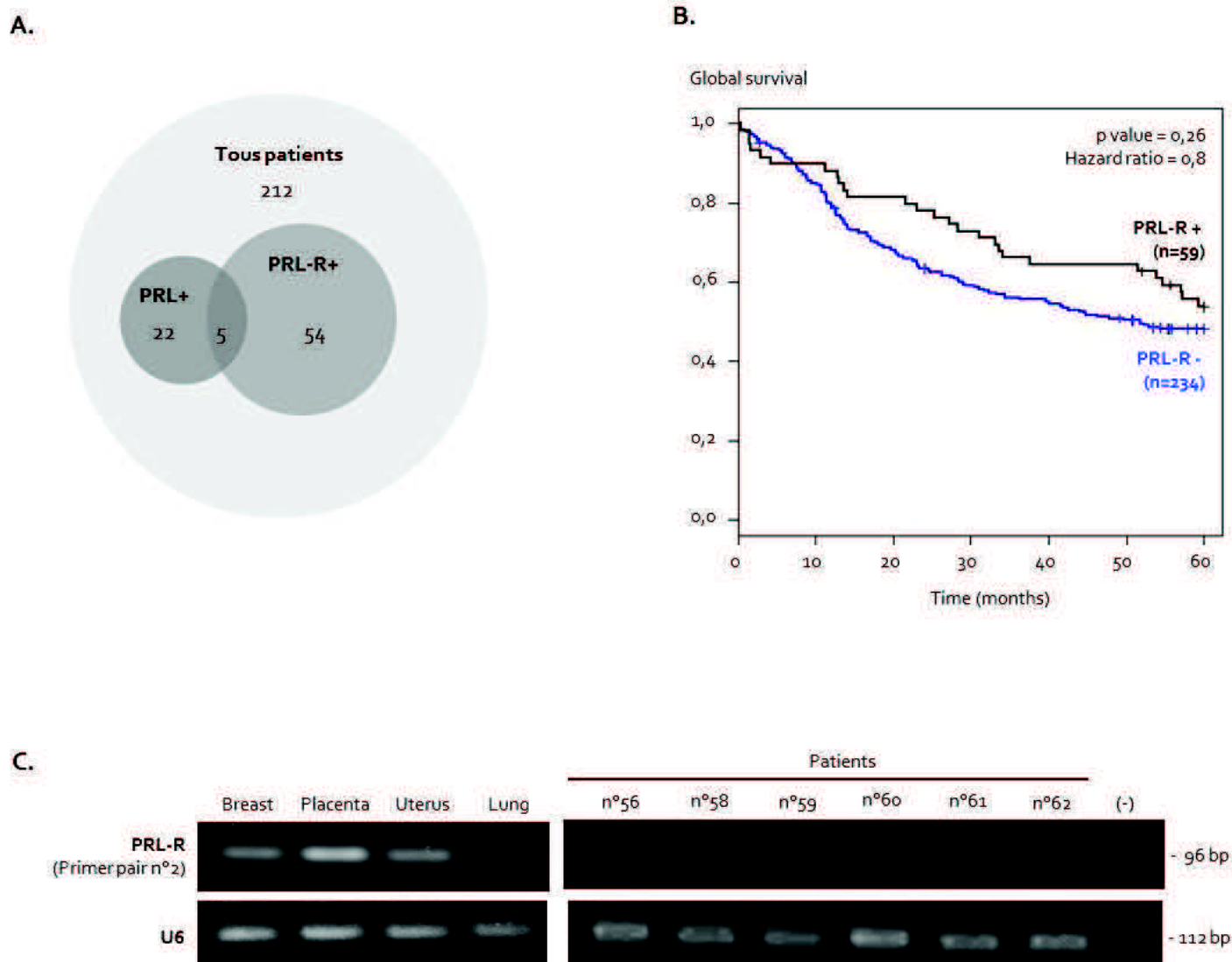


Figure 42: Expression pattern of *PRL-R* gene in patients lung tumors

A. Expression of *PRL-R* gene in lung cancer patients.

Signal ratios of the *PRL-R* probe set were obtained by the transcriptomic analysis (Affymetrix technology) of 293 patients. We consider that *PRL-R* gene is expressed within the tumor if the *PRL-R* expression ratio is higher than a threshold established by averaging the *PRL-R* signal in the 109 human somatic tissues. 56 tumors exhibiting a *PRL-R* expression ratio higher than the threshold are considered as *PRL-R* expressing tumors: patients are renamed *PRL-R* patients. This Venn diagram illustrates the *PRL* and *PRL-R* status of the 293 patients of the cohort. Only 5 patients among the 27 *PRL* patients express the *PRL-R* gene within their tumor.

B. Survival probability curves (estimated by the Kaplan-Meier method) for *PRL-R* (n=234) and *PRL-R* patients (n=56). No significant difference in prognosis is shown by the univariate logrank test.

C. 6 patient samples were submitted to PCR with specific *PRL-R* primers (recognizing all *PRL-R* mRNA variants). *U6* amplification is used as loading control. Negative control (-) is performed with H_2O .

d'entre elles, la lignée H526 (*PRL*-négative), alors qu'aucune expression ne peut être détectée dans les 3 lignées *PRL*-positives (H146, H524, H69). Suite à ce résultat inattendu, je me suis demandé si les cellules *PRL*+ ne pourraient pas exprimer une autre isoforme du *PRL*-R dont l'expression ne serait pas détectée avec la paire de primers utilisée. Un alignement des différentes isoformes protéiques du récepteur décrites à ce jour m'a permis d'identifier un sous-domaine présent dans l'ensemble des isoformes : il s'agit du second domaine extracellulaire de type fibronectine III (**Figure 41C**). On conçoit alors une nouvelle paire de primers dans la séquence nucléotidique codant ce sous-domaine commun à l'ensemble des variants *PRL*-R, en prenant soin que ces primers ne reconnaissent pas également d'autres récepteurs à cytokines chez qui la présence d'un domaine fibronectine de type III est courante. Avec cette nouvelle paire de primers, c'est à nouveau dans la seule lignée H526 que je détecte une expression de *PRL*-R en qPCR (**Figure 41D**). Ces premiers résultats semblent indiquer que le rôle de l'expression aberrante de *PRL* sur le phénotype cancéreux des cellules de cancer du poumon à petites cellules ne nécessite pas la présence de son récepteur.

Nous nous sommes ensuite penchés sur les données transcriptomiques des tumeurs de patients de la cohorte (**Figure 42**). L'établissement d'un seuil nous a permis de définir les patients *PRL*-R+ : au nombre de 54, ces patients expriment le gène *PRL*-R au sein de leur tumeur. Néanmoins, les 27 patients *PRL*+ ne sont pas tous *PRL*-R+ : seuls 5 d'entre eux expriment de manière conjointe le récepteur et son ligand (**Figure 42A**). De plus, si notre hypothèse d'une expression aberrante de la prolactine et de son récepteur agissant dans une boucle autocrine/paracrine était juste, alors l'expression du gène *PRL*-R serait également associée à un pronostic sombre. Ce n'est pas ce qui est observé (**Figure 42B**), puisqu'aucune différence significative de survie n'est observée entre les groupes de patients *PRL*-R- et *PRL*-R+. Ceci a été confirmé dans deux études externes (GSE19188 et LgKcBeer) où aucune association de l'expression *PRL*-R à la survie n'a été observée (données non présentées).

2.3.2. La prolactine exprimée de manière ectopique pourrait agir *in cellulo* sans être sécrétée

Les résultats présentés précédemment suggèrent que le récepteur à la prolactine n'est pas toujours exprimé par les cellules cancéreuses pulmonaires *PRL*-positives et que sa présence n'est pas nécessaire à l'acquisition d'une chimio-résistance accrue par les cellules qui expriment le gène *PRL* de manière aberrante. Il nous faut alors revoir notre hypothèse d'une action de la prolactine dans une boucle autocrine/paracrine impliquant la liaison de la protéine à son récepteur membranaire. Nous avons alors émis une nouvelle hypothèse, originale et inédite, selon laquelle la prolactine pourrait agir à l'intérieur de la cellule, indépendamment de son récepteur, sans être sécrétée.

J'ai tenté de détecter la protéine prolactine dans les lignées cellulaires *PRL*+ grâce à un anticorps anti-*PRL* (**Figure 43**). Je n'ai réussi à observer la présence de la protéine ni par immunofluorescence (**Figures 43C et D**), ni par western-blot, que ce soit dans les extraits cellulaires (**Figure 43A**) ou dans le milieu de culture (**Figure 43B**). Les essais de détection en immunohistochimie sur coupes de tumeurs *PRL*+ se sont également révélés vains (données non présentées). Au-delà de l'idée selon laquelle l'ARN messager peut ne pas être traduit en protéine, nous avons pensé dans un premier temps que la prolactine pourrait être exprimée faiblement ou être peu stable, et donc non détectable par cet anticorps dont le fonctionnement n'a été vérifié que dans l'hypophyse qui exprime la prolactine à un niveau très élevé, les cellules cancéreuses de prostate qui surexpriment fortement la prolactine ([Dagvadorj et al., 2007](#)) et sur prolactine recombinante purifiée (**Figure 43A et B**). Je suis donc restée, jusqu'à aujourd'hui, dans l'incapacité de démontrer la sécrétion de la protéine dans le milieu extérieur, ou *a contrario* sa rétention dans les cellules cancéreuses pulmonaires.

L'absence de récepteur à la prolactine nous a conduit à imaginer que la prolactine ne soit pas sécrétée par les cellules cancéreuses pulmonaires, et à chercher des explications possibles à cette absence de sécrétion. Dans un premier temps, nous avons souhaité vérifier l'intégrité de la machinerie de sécrétion dans les cellules

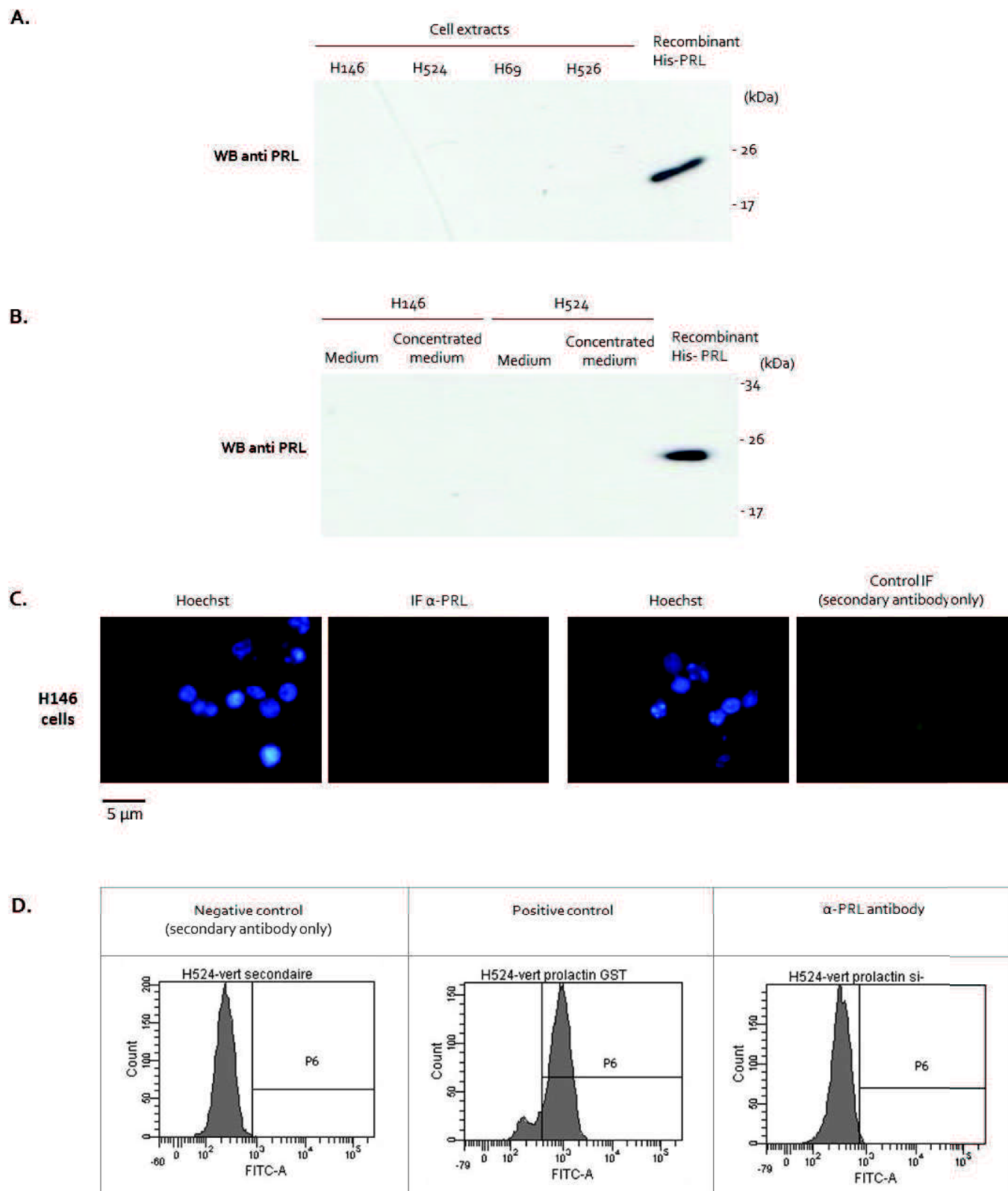


Figure 43: PRL protein is neither detected in *PRL*+ SCLC cell lines extracts, nor in cell culture supernatants

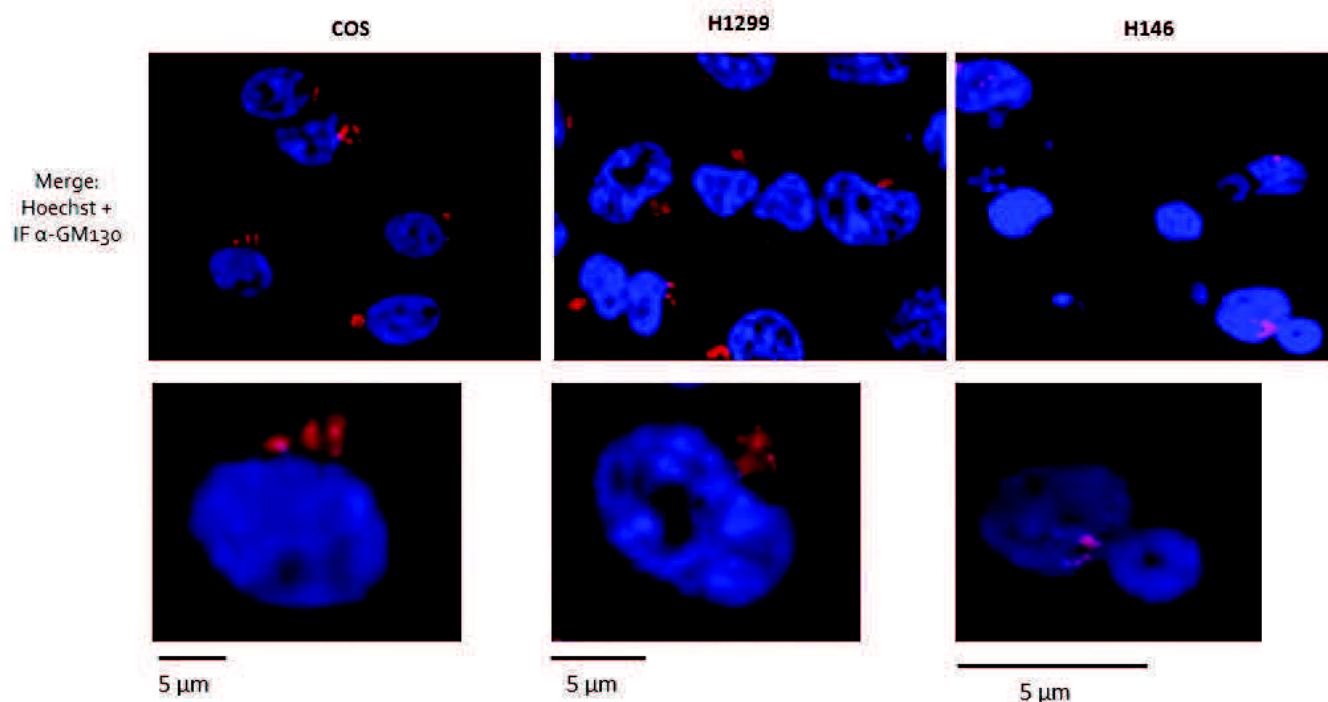
A. Detection of PRL protein in SCLC cell extracts by SDS-PAGE with anti-PRL antibody (BioGenex, 1/400). Recombinant purified His-PRL is used as positive control.

B. Detection of PRL protein in SCLC cells culture supernatant (medium) or culture supernatant concentrated 10 times by TCA precipitation (concentrated medium) by SDS-PAGE with anti-PRL antibody (BioGenex, 1/400). Recombinant purified His-PRL is used as positive control.

C. Detection of PRL protein on fixed SCLC H146 cells by immunofluorescence (IF) with anti-PRL antibody (BioGenex, 1/40). Non-specific staining is evaluated with anti-mouse 488 Alexa secondary antibody alone.

D. Quantification of immunofluorescence on SCLC H524 cells by flow cytometry. Positive control is obtained by incubating cells with recombinant GST-PRL.

A.



B.

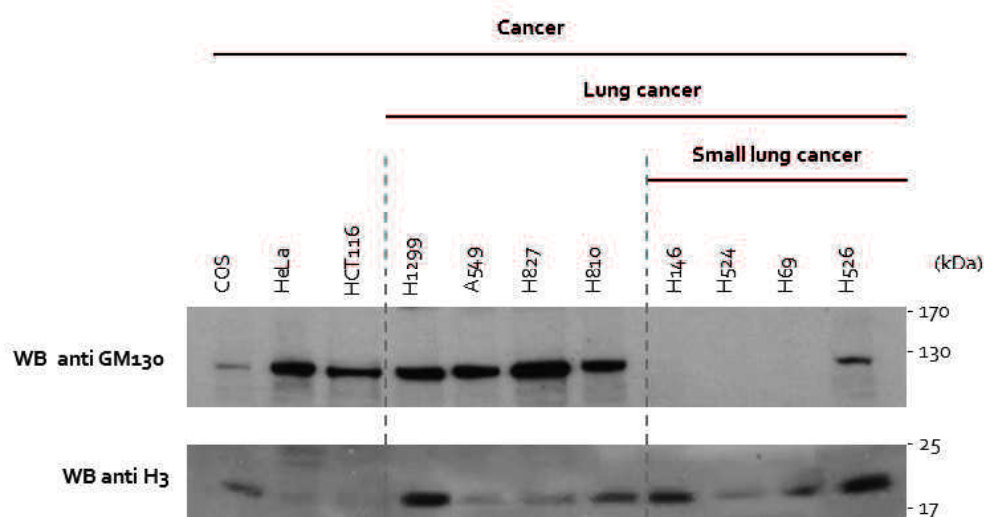
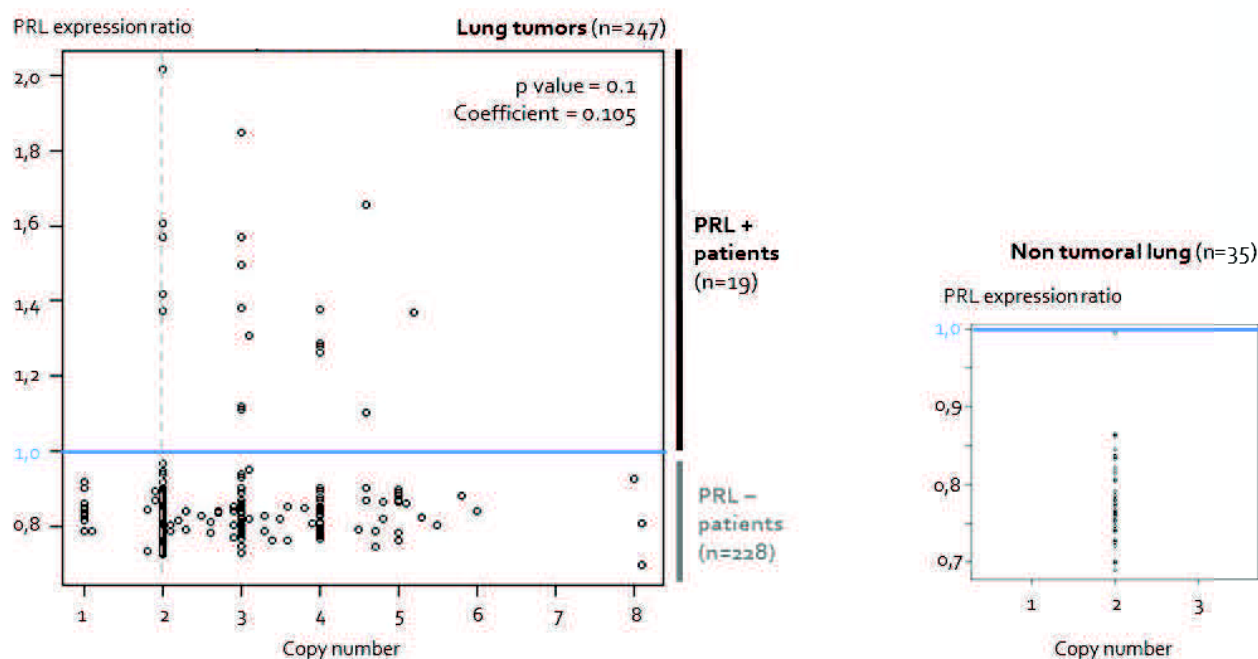


Figure 44: An impaired secretion of prolactin due to a low functional Golgi apparatus?

A. Detection of the Golgi structural protein GM130 on fixed cancer cells (COS, H1299 and the SCLC H146 cells) by immunofluorescence.

B. Detection of GM130 protein in cancer cell extracts by SDS-PAGE. H3 staining is used as a loading control.

A.



B.

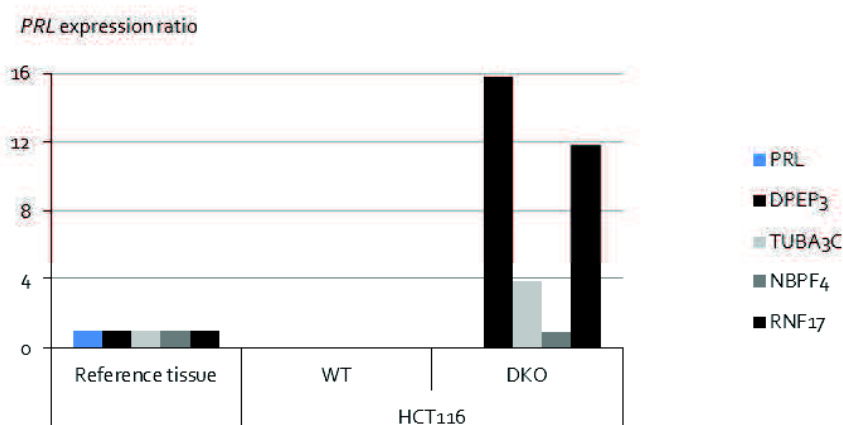


Figure 45: Mechanisms of *PRL* gene re-expression in lung cancer cells

A. Analysis of variations in copy number (deletion and amplification) at the *PRL* gene locus by CGH array in 247 lung tumors of the cohort (left panel). *PRL* expression ratio is indicated on the y-axis. The correlation between *PRL* expression and *PRL* copy number is evaluated by the non-parametric Spearman's coefficient. Control analysis is performed on the paired non-tumoral lung samples from 35 patients (right panel).

B. Expressions of *PRL*, as well as 4 testis-specific genes (*DPEP3*, *TUBA3C*, *NBPF4* and *RNF17*) were analysed by RT-qPCR in the HCT116 cell line, wild-type and mutant (DKO: double knock-out for *DNMT1* and *DNMT3b*). Expressions are normalized thanks to control genes. A value of 1 is assigned to the expression in the reference tissue (placenta for *PRL*, testis for other genes).

cancéreuses pulmonaires, et plus particulièrement l'intégrité de l'appareil de Golgi (**Figure 44**). Dans les cellules cancéreuses simiennes COS et dans les cellules humaines de cancer du poumon non à petites cellules *PRL*-négatives H1299, l'appareil de Golgi est très facilement détecté en immunofluorescence grâce à un anticorps reconnaissant spécifiquement la protéine golgienne GM130 (**Figure 44A**). L'appareil du Golgi y apparaît comme une structure unique dans la cellule, bien délimitée et apposée contre le noyau. Le marquage est ponctiforme, signant la présence de plusieurs saccules golgiennes empilées. A l'inverse, le signal est beaucoup plus faible et plus diffus dans les cellules de CBPC *PRL*-positives H146. Dans certaines cellules, aucun signal anti-GM130 ne peut être détecté, alors que dans d'autres il colocalise dans le noyau avec le Hoechst. On a ensuite cherché à détecter cette protéine golgienne en western-blot et à quantifier son abondance dans divers types de cellules cancéreuses (**Figure 44B**). Alors que la protéine est toujours détectée à la taille attendue (130 kDa), bien qu'à des niveaux variables, dans les autres types de cellules cancéreuses testés (incluant des cellules CBNPC) et dans la lignée CBPC *PRL*-négative H526, aucun signal n'est obtenu dans les extraits de cellules CBPC *PRL*-positives H146, H524 et H69. Ces observations suggèrent que l'appareil de Golgi est particulièrement réduit, voire désorganisé, dans les cellules CBPC *PRL*+. Ceci peut être appuyé par le fait que les cellules de cancer du poumon à petites cellules sont, comme leur nom l'indique, des cellules de petite taille, à cytoplasme restreint et à faible contenu en organites. En microscopie électronique également, l'appareil de Golgi est difficilement détectable dans les cellules CBPC et ne prend jamais l'aspect classique en saccules empilées observable dans les cellules CBNPC. Néanmoins, si on peut imaginer une faible fonctionnalité de l'appareil de Golgi dans ces cellules, il est difficile de croire que la fonction golgienne soit complètement abolie, étant donnés les rôles assurés par l'appareil de Golgi dans les cellules eucaryotes (modifications post-traductionnelles de protéines, bourgeonnement des vésicules pour l'exocytose, envoi à la membrane de nombreux récepteurs membranaires – et dans le cas des cancers du poumon à petites cellules, sécrétion de nombreux peptides neuroendocrines). Ainsi, la faible fonctionnalité du Golgi dans les cellules *PRL*+, même si elle était avérée, ne suffirait à elle seule à expliquer l'absence de sécrétion de la prolactine. Mais nous allons le voir, c'est sur une étude plus poussée des modes d'expression du gène *PRL* dans les cellules cancéreuses pulmonaires que nous avons pu appuyer notre hypothèse d'une prolactine non sécrétée.

2.4. L'expression aberrante du gène *PRL* dans les cellules pulmonaires cancéreuses aboutit à la production d'un ARN messager tronqué en 5'

2.4.1. L'expression ectopique du gène *PRL* ne semble due ni à une amplification génique, ni à une déméthylation du promoteur

Suite à notre observation inédite de l'expression du gène *PRL* dans le poumon tumoral, nous avons tenté d'émettre des hypothèses sur la cause de cette ré-expression aberrante d'un gène normalement réprimé dans le poumon sain. La première hypothèse que nous avons envisagée est que le gène *PRL* ait pu subir durant la carcinogenèse pulmonaire une amplification de sa séquence, du fait d'événements de duplication ou de recombinaison non homologue par exemple, conduisant à son expression par les cellules cancéreuses bronchiques. Une analyse CGH array sur les ADN de tumeurs de la cohorte nous a permis d'évaluer le nombre de copies de la séquence génique *PRL* présentes dans le génome des cellules tumorales (**Figure 45A**). Ce nombre est normalement de 2, étant donné qu'un allèle est présent sur chaque chromosome de la paire 6. C'est bien ce que l'on observe dans les échantillons d'ADN de poumon non-tumoral prélevés chez ces mêmes patients (**Figure 45A, panneau de droite**). Au sein des tumeurs pulmonaires, on voit que le nombre de copies est fréquemment modifié, depuis des pertes d'homozygoties où seul l'un des deux allèles est conservé, jusqu'à des amplifications massives où la séquence *PRL* est détectée en 8 exemplaires. Pourtant, il n'existe pas de corrélation claire, mesurée par le coefficient de Spearmann, entre le nombre de copies de *PRL* sur le génome et l'expression du gène au sein de la tumeur. Ainsi, certains patients considérés comme *PRL*-négatifs possèdent une amplification massive du gène *PRL*, et à l'inverse 5 patients *PRL*-positifs ont une expression



Figure 46: The synthesis of *PRL* mRNA relies neither on the proximal pituitary promoter, nor on the distal extra-pituitary promoter

A. Primers 13 and 16 are located within the *PRL* coding sequence (CD). Primers 1 and 2 are located within the 5'UTR sequence specific of the pituitary *PRL* mRNA (variant 1). Primers 4 and 5 are located in the exon 1a sequence, that is to say in the 5'UTR sequence specific of the extra-pituitary *PRL* mRNA (variant 2).

B. Human tissues and human cancer lung cells lines RNAs were submitted to RT. PCR analysis was performed using RT products as templates, with the various previously described primers. *U6* amplification is used as loading control. Negative control (-) is performed with H₂O.

aberrante du gène alors que la séquence génique n'est présente qu'en ses deux copies originelles. Bien que les 22 autres patients *PRL*⁺ présentent un nombre de copies supérieur à 2, cette amplification du gène *PRL* ne semble pas pouvoir expliquer à elle seule son expression ectopique par les cellules tumorales.

La seconde hypothèse est que le gène *PRL* ait pu subir une déméthylation aberrante conduisant à sa ré-expression par les cellules cancéreuses. J'ai testé cette hypothèse de manière indirecte, en utilisant une lignée cellulaire cancéreuse ne présentant plus d'activité DNA méthyltransférase (**Figure 45B**). Cette lignée HCT116 double KO (DKO) pour DNMT1 et DNMT3b présente une hypométhylation drastique globale du génome, puisqu'on estime que 95% de la méthylation est perdue, et ré-exprime de manière aberrante plusieurs gènes initialement réprimés, dont le gène suppresseur de tumeur P16INK4a et certains gènes spécifiques de la lignée germinale mâle (Rhee et al., 2002). Le gène *PRL* n'est pas exprimé dans la lignée HCT116 sauvage (**Figure 36A**) et aucune ré-expression du gène *PRL* n'est observée par qPCR dans la lignée DKO (**Figure 45B**), contrairement à 3 gènes testiculaires testés (DPEP3, TUBA3C et RNFI7). Ce résultat suggère qu'une déméthylation n'est pas suffisante pour obtenir une ré-expression du gène *PRL* dans les cellules cancéreuses.

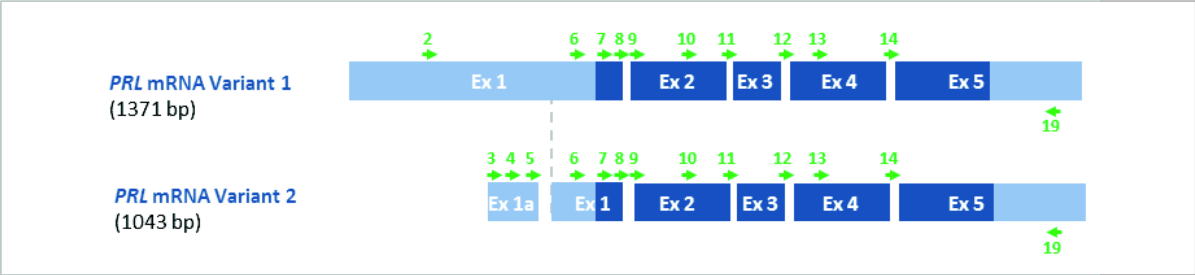
2.4.2. L'expression du gène *PRL* dans les cellules cancéreuses pulmonaires ne repose sur l'utilisation d'aucun des deux promoteurs connus de *PRL*

Ni une amplification génique, ni une déméthylation du promoteur ne semblent pouvoir expliquer à elles-seules la ré-expression ectopique du gène *PRL* dans les cellules cancéreuses pulmonaires. Nous avons alors voulu savoir lequel des promoteurs du gène *PRL* était utilisé dans le poumon tumoral. L'expression du gène *PRL* est connue pour être régulée de manière tissu-spécifique grâce à la présence de deux promoteurs distincts : un promoteur proximal hypophysaire permettant la synthèse d'un variant 1 d'ARN messager contenant la séquence de l'exon 1 en entier, et un second promoteur plus distal utilisé dans les sites extra-hypophysaires de production de la prolactine, conduisant à la synthèse d'un variant 2 qui inclue la séquence d'un exon alternatif 1a et seulement la partie 5' de l'exon 1 (voir Introduction, paragraphe 1.1.3 et **Figure 1**). J'ai tenté de savoir lequel de ces deux promoteurs dirige l'expression hors-contexte du gène *PRL* dans les cellules CBPC (**Figure 46**) en adoptant une approche par PCR, grâce à des primers forward reconnaissant les parties 5'UTR de l'ARN messager propres à chacun des variants (**Figure 46A**). L'utilisation des paires de primers spécifiques du variant 1 conduit bien à une amplification dans l'hypophyse, et une bande est détectée dans le placenta avec les primers spécifiques du variant 2 (**Figure 46B**). De manière inattendue, aucune amplification n'est obtenue dans les lignées CBPC *PRL*⁺. Nous en concluons qu'aucun des deux promoteurs connus du gène *PRL* n'est utilisé dans les cellules *PRL*⁺, et qu'un nouveau variant d'ARN messager, n'incluant pas les séquences reconnues par les primers, est synthétisé.

2.4.3. Un ARN messager tronqué ne contenant pas la séquence de l'exon 1 est produit dans les lignées de cancer du poumon à petites cellules

Suite à cette observation inattendue, j'ai procédé à une cartographie des ARN messagers *PRL* produits par les cellules CBPC *PRL*-positives (**Figure 47**). Ceci a été réalisé, toujours par PCR, en utilisant un primer reverse unique localisé dans la partie 3'UTR, et en balayant la séquence *PRL* avec une succession de primers localisés toujours plus en amont dans chacun des exons (**Figure 47A**). Dans le placenta, on détecte une amplification avec l'ensemble des primers, à l'exception du primer situé dans la partie 5'UTR propre au variant hypophysaire (**Figure 47B**). Dans les lignées CBPC H146 et H524, les amplifications sont correctes lorsque l'on utilise les primers situés en aval de l'exon 1 (jusqu'au primer n°9 situé à cheval sur les séquences des exons 1 et 2). Puis, à partir du primer n°8, qui reconnaît les 20 derniers nucléotides de l'exon 1, plus aucune amplification n'est obtenue. Dans la lignée H69, les observations sont identiques, si ce n'est que le primer n°9 ne fournit pas non plus une amplification correcte avec une bande unique à la taille attendue de 709 pb. Ainsi, les séquences des exons 1 et 1a sont absentes des transcrits *PRL* produits par les cellules CBPC *PRL*-positives.

A.



B.

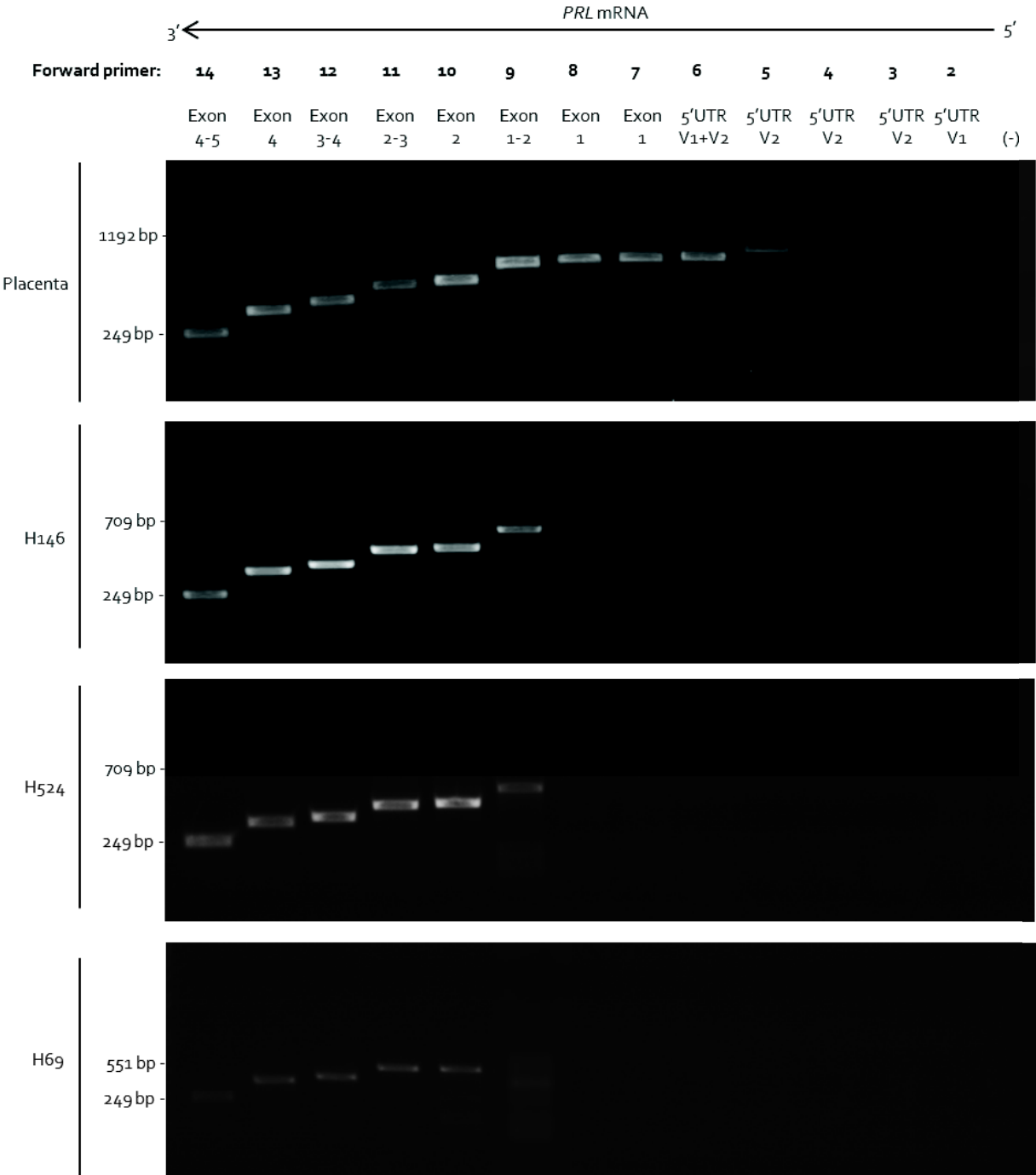
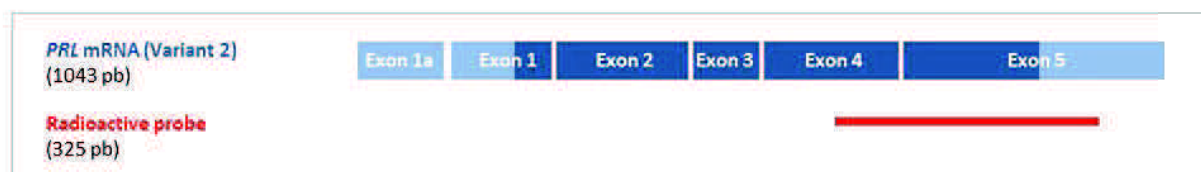


Figure 47: Cartography of *PRL* mRNA reveals that exon 1 and exon 1a sequences are absent of the SCLC cell lines transcripts

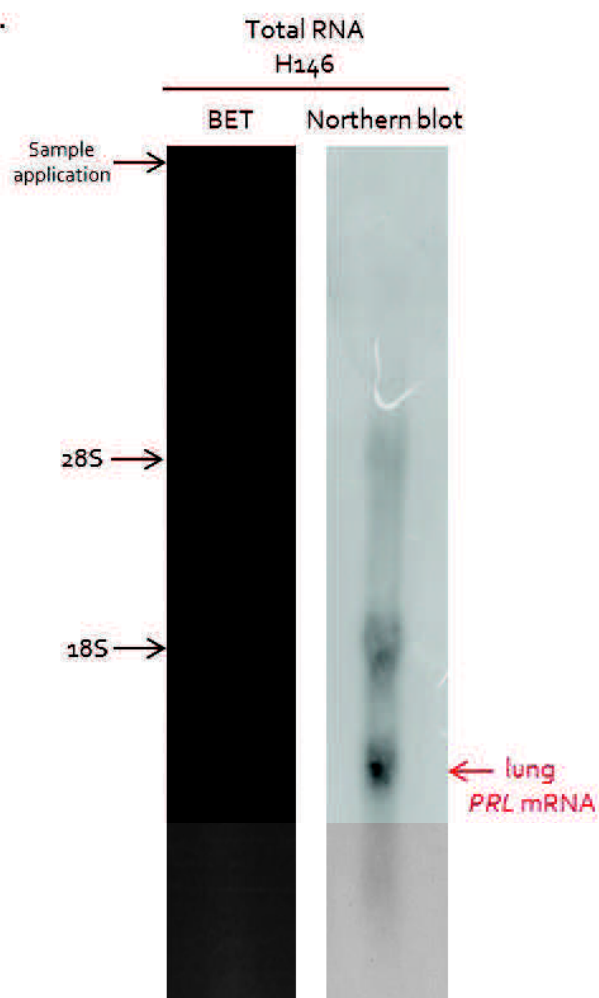
A. Cartography of the *PRL* mRNA produced in SCLC *PRL*-positive cell lines was performed by using a reverse primer in 3'UTR and various forward primers along the *PRL* transcript sequences (coding sequence and 5'UTR of the two distinct variants).

B. *PRL*-positive SCLC cells lines RNAs were submitted to RT. PCR analysis was performed using RT products as templates, with the various primers. Negative control (-) is performed with H₂O.

A.



B.



C.

Observed band	Expected size	Distance	Calculated size
28 s	4,7 kb	3,96 cm	
18 s	2,3 kb	6,52 cm	
Signal		8,09 cm	828 pb

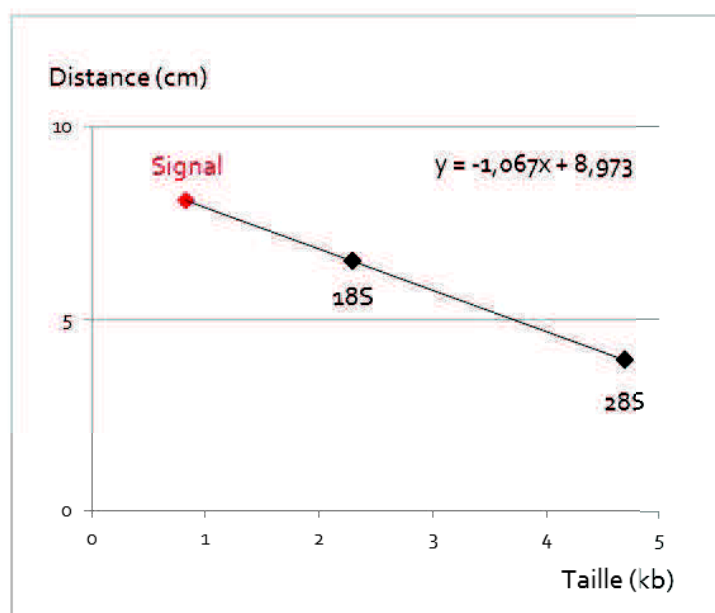


Figure 48: Analysis of the *PRL* mRNA produced by H146 SCLC *PRL*+ cells by Northern blot

A. The radioactive synthesized probe specifically recognizes exons 4 and 5 of *PRL* gene sequence.

B. After migration of total RNAs in an agarose gel, BET stained ribosomal RNAs (28S and 18S) are visible under UV light (left panel). After transfert and hybridization of the probe on the membrane, a film is exposed and signal is revealed by autoradiography (right panel).

C. Calculation of the size of the visible band is made using the known sizes of ribosomal RNAs and distances.

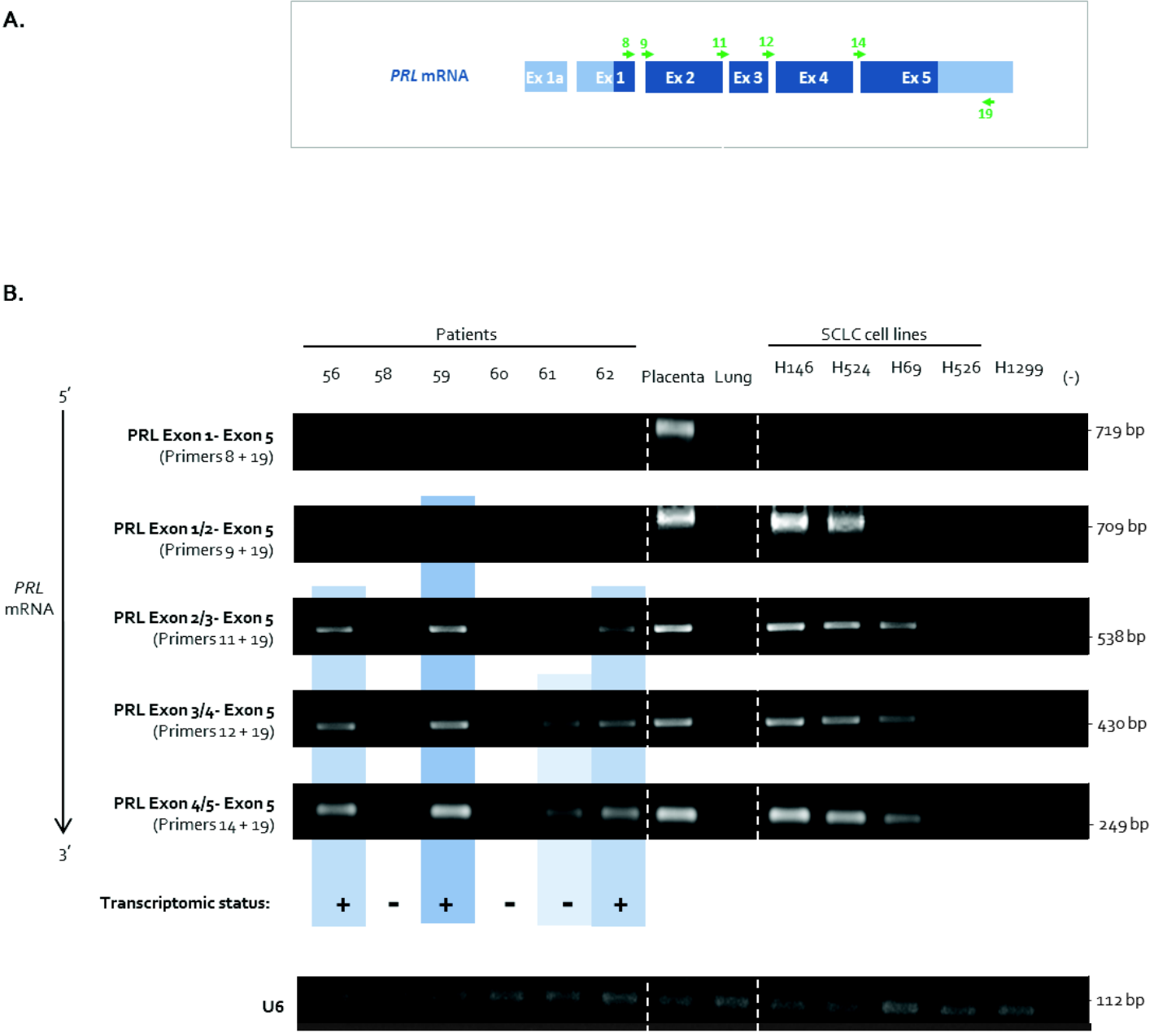


Figure 49: Cartography of *PRL* mRNA in tumor samples reveals differently extended absences of 5' sequences and reveals new *PRL*+ patients that were not detected by the transcriptomic analysis

A. Cartography of the *PRL* mRNAs produced in patients lung tumors was performed by using a reverse primer in 3'UTR and 5 differents forward primers corresponding to every exons junction sequences.

B. Tumor RNAs were submitted to RT. PCR analysis was performed using RT products as templates, with the various primers. *U6* amplification is used as loading control. Negative control (-) is performed with H₂O. The *PRL* status (+: expressing or -: not expressing) as determined by the transcriptomic analysis is indicated for each patient. Blue rectangles highlight observed amplifications in patients.

J'ai réalisé un Northern blot sur ARN totaux de cellules CBPC. Ceci nous permet d'observer la taille de l'ARN messager mature *PRL* produit par ces cellules cancéreuses pulmonaires (**Figure 48**). Le gel d'électrophorèse des ARN, photographié sous UV avant transfert, me permet d'observer les ARN ribosomiaux 28S (4,7 kb) et 18S (2,3 kb) (**Figure 48B**). Après transfert et hybridation de la membrane avec une sonde radioactive spécifique des exons 4 et 5 du gène *PRL* (**Figure 48A**), on observe le marquage non spécifique des ARN ribosomiaux, mais également un signal pouvant correspondre à l'ARNm *PRL* (**Figure 48B**) qui n'est pas détecté lorsque les ARN des cellules H526 *PRL*-négatives sont analysés (contrôle négatif, non présenté). Un calcul basé sur les distances parcourues sur le gel et les tailles connues des ARN ribosomiaux me permet de calculer une taille théorique de 800 pb pour le signal obtenu (**Figure 48C**), taille compatible avec un ARN messager privé de ses séquences exoniques 1a et 1.

2.4.4. Les ARN messagers synthétisés au sein des tumeurs pulmonaires de patients ne contiennent pas les séquences exoniques en 5'

J'ai réalisé une analyse basée sur le même principe sur les échantillons issus de tumeurs de patients (**Figure 49**). Le même primer reverse dans le 3'UTR est utilisé, et seuls 5 primers forward, localisés aux jonctions des exons, sont utilisés sur ces échantillons précieux (**Figure 49A**). Lorsque l'on cherche à amplifier depuis l'intérieur de l'exon 1 (primer n°8), seul le transcrit placentaire est détecté (**Figure 48B**). Avec le primer n°9 situé à cheval sur les exons 1 et 2, on détecte des amplifications dans nos lignées H146 et H524, mais également une faible amplification chez le patient n°59 qui était bien considéré comme *PRL*+ en transcriptomique. Il faut amplifier depuis la jonction entre les exons 2 et 3 pour détecter la lignée H69, ainsi que les deux autres patients *PRL*+ en transcriptomique (n°56 et 62). De manière inattendue, une amplification encore plus en aval, avec un primer situé au début de l'exon 4, est détectée chez le patient n°61, patient qui avait pourtant été considéré comme *PRL*-négatif lors des analyses transcriptomiques (**Figure 35B**). Nous en déduisons que les patients *PRL*+ expriment, tout comme les lignées cellulaires SCLC, des ARN messagers *PRL* dans lesquels les séquences en 5' sont manquantes. L'étendue de la zone manquante s'avère variable, depuis les séquences des exons 1 et 1a uniquement jusqu'à l'absence des 2 premiers exons entiers, voire davantage. De plus, ce résultat nous enseigne que certains patients, qui avaient été considérés comme *PRL*-négatifs suite aux analyses transcriptomiques, s'avèrent en fait exprimer de manière aberrante le gène *PRL*, même si le transcrit synthétisé est alors privé d'une grande partie de ces séquences en 5', et qu'il n'aurait pas dû être détecté de manière efficace par le probeset *PRL* utilisé. Il faut donc envisager qu'aux 10% de patients *PRL*+ déjà identifiés s'en ajoutent de nouveaux dont seule la détection du transcrit en PCR pourrait démontrer le statut *PRL*+

Nous avons démontré que les cellules cancéreuses pulmonaires produisent un ARN messager *PRL* privé des séquences des exons 1a et 1. Or ces exons correspondent non seulement à la partie 5'UTR non traduite, mais également au début de la séquence codante qui formera après traduction le peptide signal de la prolactine. Ces résultats obtenus sur l'ARN messager tendent donc à conforter notre hypothèse de la production par les cellules cancéreuses pulmonaires d'une prolactine non sécrétable.

2.5. Quelques données sur l'action *in cellulo* de la prolactine dans les cellules cancéreuses pulmonaires

2.5.1. Mise au point d'une lignée modèle de cancer du poumon exprimant une prolactine non sécrétable

Pour tenter de décrypter les mécanismes qui pourraient être initiés par la prolactine retenue au sein des cellules cancéreuses pulmonaires, nous avons mis au point un modèle cellulaire de prolactine non-sécrétable

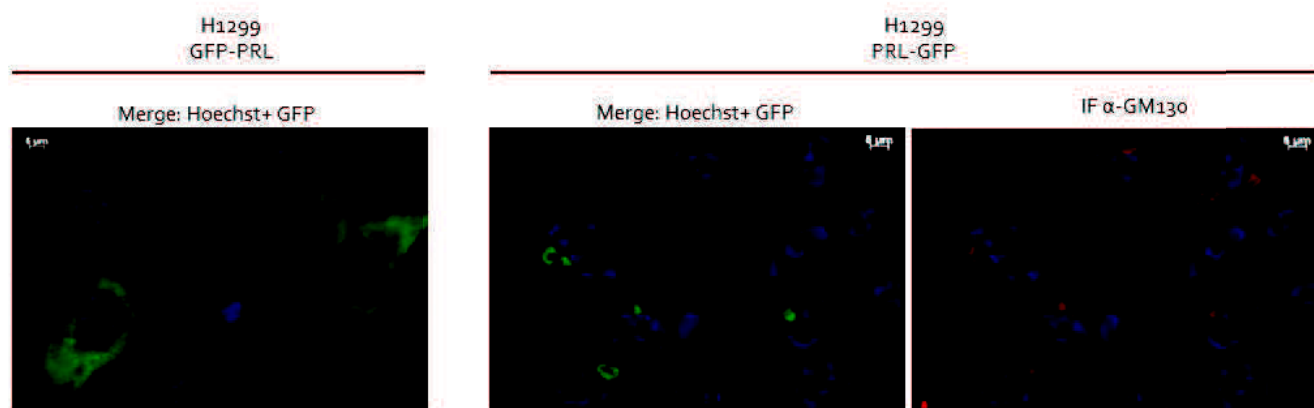
A.



B.



C.



D.

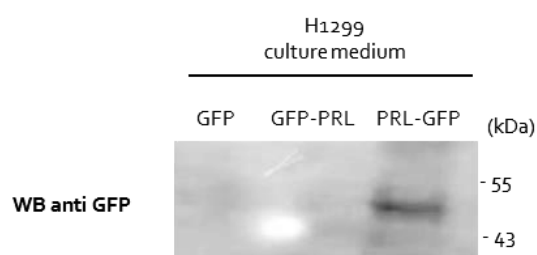


Figure 50: Expression of exogenous secretable PRL-GFP and non-secretable GFP-PRL in the H1299 human lung cancer cell line

A. Detection of the exogenous fusion *PRL* gene in H1299 cells transfected with peGFP-C-PRL and peGFC-N-PRL plasmids. H1299 WT cells and H1299 cells transfected with empty vector are used as control.

B. Detection by SDS-PAGE with anti-GFP antibody (Covance, 1/5000) of the GFP-tagged forms of PRL expressed in H1299 cells.

C. Detection of the fluorescence of GFP-tagged forms of PRL expressed in H1299 cells. Nuclei are labeled with Hoechst. Co-immunodetection of the Golgi structural protein GM130 is performed in H1299 cells expressing the PRL-GFP fusion protein.

D. Detection of GFP-tagged forms of PRL in H1299 cells culture supernatants concentrated 10 times by TCA precipitation by SDS-PAGE with anti-GFP antibody.

(**Figure 50**). Des vecteurs codants des formes de la prolactine fusionnées à la GFP du côté amino-terminal (GFP-PRL) ou carboxy-terminal (PRL-GFP) ont été transfectés dans des cellules H1299 de cancer du poumon non à petites cellules. Puis des lignées exprimant ces constructions de manière stable ont été obtenues par sélection avec un antibiotique et tri cellulaire en cytométrie de flux. L'absence d'expression du gène *PRL* endogène dans la lignée mère (WT) et dans la lignée contrôle transfectée avec le vecteur vide (GFP) a été vérifiée en PCR (**Figure 50A**). D'autre part, on amplifie bien les séquences transfectées codant la prolactine dans nos deux lignées d'intérêt (GFP-PRL et PRL-GFP). La présence des protéines de fusion est avérée en western blot grâce à un anticorps reconnaissant la GFP (**Figure 50B**). Ces protéines ont des masses apparentes de 55 kDa environ, reflet de la fusion de la prolactine (25 kDa) avec la GFP (30 kDa). Puis les localisations subcellulaires des protéines de fusion sont observées en microscopie à fluorescence (**Figure 50C**). La protéine de fusion GFP-PRL exhibe une localisation majoritairement cytoplasmique dans les cellules qui l'expriment, avec un léger marquage dans le noyau. A l'inverse, la protéine PRL-GFP exhibe un signal ponctué, avec un spot unique dans chaque cellule qui l'exprime. Cette PRL-GFP colocalise parfaitement avec la protéine golgienne GM130. Enfin, j'ai cherché à détecter les protéines de fusion dans le milieu de culture des cellules (**Figure 50D**). La GFP et la protéine GFP-PRL ne sont pas détectées dans le surnageant des cellules, même concentré. Seule la présence de la protéine PRL-GFP peut être observée dans le milieu. Nous concluons de ces résultats que nous disposons de deux modèles cellulaires dans lesquels les comportements de la protéine prolactine sont très différents. Dans les cellules H1299 PRL-GFP, la prolactine emprunte la voie de sécrétion en transitant notamment par l'appareil de Golgi, puis est sécrétée dans le milieu extérieur. Elle adopte ainsi le comportement classique de la prolactine avec un peptide signal fonctionnel, comme ce qui est observé dans les cellules lactotrophes de l'hypophyse ou dans les cellules de la prostate. A l'inverse, dans la lignée H1299 GFP-PRL, la prolactine produite n'est pas sécrétée dans le milieu extérieur, elle est retenue dans le cytoplasme de la cellule. Le peptide signal de la prolactine doit ne pas pouvoir assurer son rôle consistant à orienter la protéine de fusion vers la voie de sécrétion. Nous pensons que cette lignée peut être un bon modèle pour étudier le comportement d'une prolactine non sécrétable, retenue dans une cellule cancéreuse pulmonaire.

Avant toute chose, j'ai voulu savoir si cette expression d'une prolactine non sécrétable dans la lignée H1299 était capable d'induire le phénotype de chimio-résistance observé dans les lignées CBPC *PRL*-positives. J'ai réitéré le test de mortalité induite par l'Actinomycine D en cytométrie de flux (**Figure 51A**). Dans un premier temps (**panneaux de gauche**), nous avons observé que les 3 lignées établies (H1299 GFP, GFP-PRL et PRL-GFP) étaient plus sensibles à la mort cellulaire induite par un traitement à l'actinomycine D que la lignée-mère H1299. Ceci va à l'encontre de l'effet chimio-protecteur que nous attendions. Mais il faut noter qu'au sein d'une lignée stable, et malgré la sélection appliquée, l'ensemble des cellules n'expriment pas le transgène. J'ai alors, dans l'analyse des données de cytométrie, distingué les cellules non fluorescentes qui n'expriment pas le transgène, et les cellules dont la fluorescence de GFP, GFP-PRL ou PRL-GFP est détectée par le FACS (**panneaux de droite**). Cette analyse tend à démontrer que ce ne sont pas uniquement les cellules qui expriment les transgènes *PRL* qui sont plus sensibles à l'apoptose, mais bien l'ensemble des cellules des lignées stables, y compris celles qui expriment la GFP seule ou qui n'expriment pas le transgène. Ceci pourrait être dû à la sélection par un antibiotique, qui a conduit à un important effet clonal et qui a pu modifier les caractéristiques des cellules. Ainsi, ces résultats de cytométrie ne sont pas facilement interprétables et ne permettent pas d'observer l'induction d'une chimio-résistance accrue par l'expression d'une prolactine non sécrétable. J'ai également observé la prolifération des lignées H1299 GFP et H1299 GFP-PRL par simple comptage cellulaire (**Figure 51B**) et par quantification du signal Hoechst en cytométrie (**Figure 51C**). Comme dans le cas des lignées CBPC, aucune différence due à la présence de la prolactine n'a pu être observée.

2.5.2. La prolactine non sécrétable est clivée dans notre lignée modèle de cancer du poumon

Malgré l'impossibilité de reproduire le phénotype résistant des cellules dans notre lignée modèle H1299 GFP-PRL, j'ai utilisé cette lignée pour étudier le comportement intracellulaire et les caractéristiques qu'une

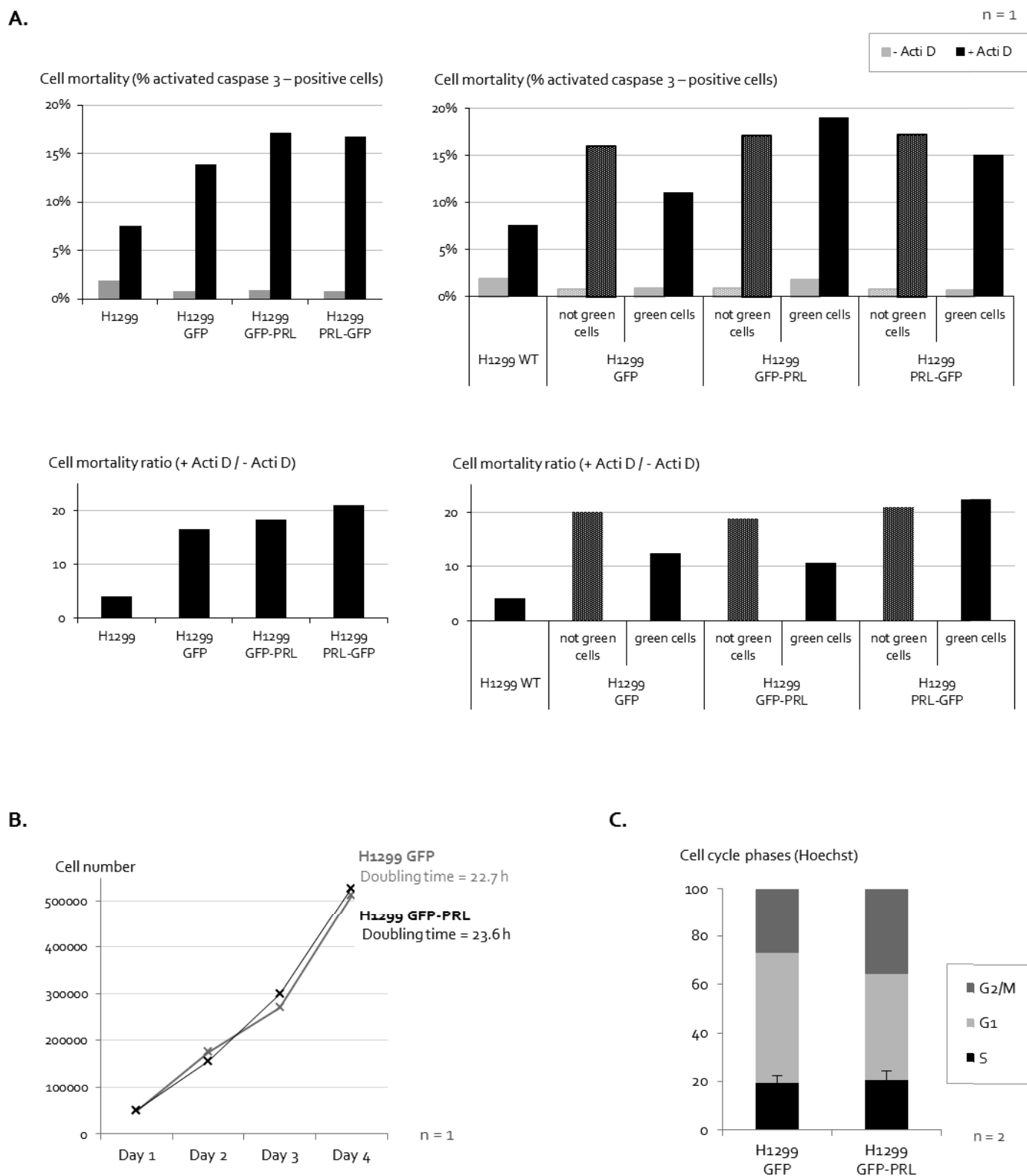


Figure 51: Effects of non-secretable GFP-PRL expression in the H1299 human cancer cell line

A. H1299 cell lines were treated for 12 hours by 250 mM Actinomycin D (Acti D) or control treatment. (Top panel) Cell mortality was measured by flow cytometry after incubation with an anti-activated Caspase 3 antibody. (Bottom panel) Cell mortality ratio was calculated by dividing the Acti D-induced cell death by the basal cell death in control condition. (Right panels) Green cells (expressing GFP or GFP fusion proteins) and cells displaying no GFP fluorescence were analysed separately.

B. Proliferation was tested by counting cells during four days.

C. The cell cycle phase of each cell line was assessed by flow cytometry with an Hoechst staining of DNA.

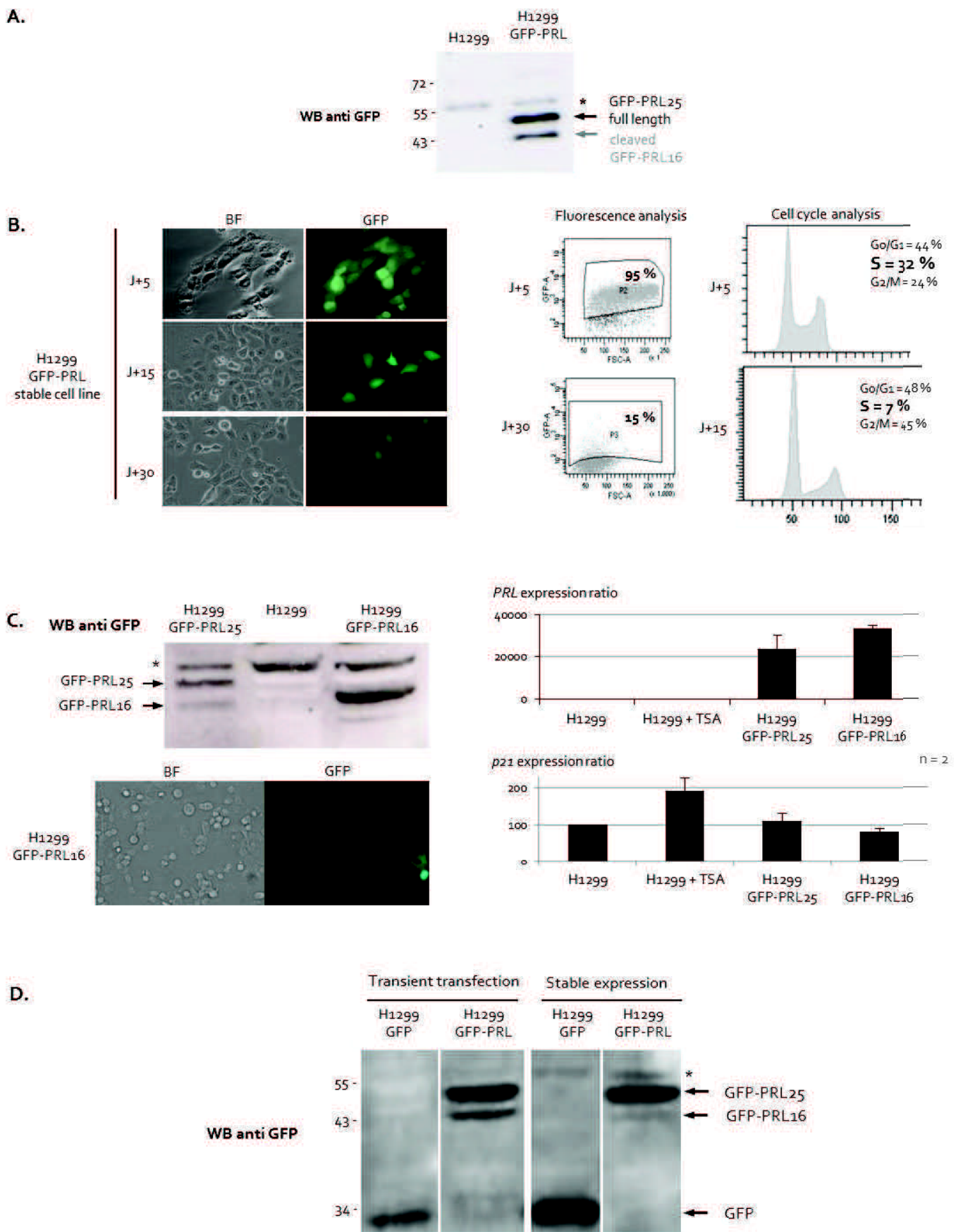


Figure 52: Non secretable GFP-PRL is cleaved in a truncated PRL16 form with potential antiproliferative role

- A.** Two distinct bands are observed in H1299 GFP-PRL extract with an anti-GFP antibody. *: non-specific band.
- B.** (Left panel) Fluorescent cells of the H1299 GFP-PRL stable cell line disappeared with time. (Right panel) Proportions of GFP-fluorescent and S-phase cells are measured by flow cytometry in the H1299 stable cell line.
- C.** (Left panel) H1299 cells are transfected with a GFP-PRL16 construction. The short form is observed with the anti-GFP antibody. (Right panel) *PRL* transgenes expression and *p21* activation are measured by qPCR. Data are normalized with *Actin*.
- D.** Abundance of the cleaved GFP-PRL16 form in GFP-PRL cell lines (in transient transfection or stable expression).

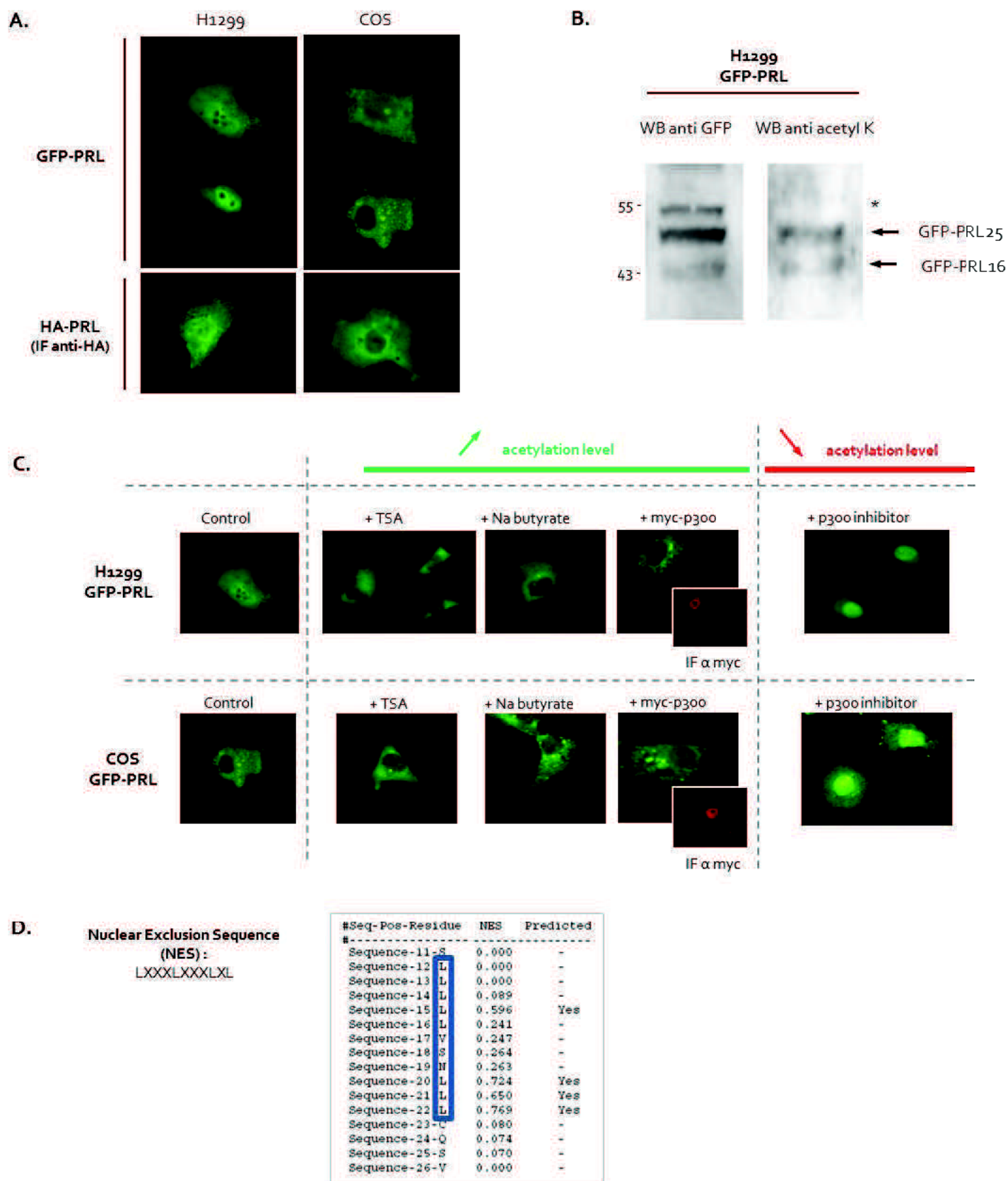


Figure 53: Acetylation of non secretable GFP-PRL controls its subcellular localization

A. Subcellular localization of GFP-PRL or HA-PRL (detected by immunofluorescence with an anti-HA antibody) differs in H1299 and COS cells.

B. The two bands corresponding to the full-length GFP-PRL and the cleaved form are detected with an anti-pan acetyl lysine antibody.

*: non-specific band.

C. Subcellular localization of GFP-PRL in H1299 or COS cells varies when the acetylation level is modulated by acetylation drugs or by the co-transfection with the acetyl-transferase p300 (detected in immunofluorescence with an anti-myc antibody).

D. The server « NetNES 1.1 Prediction » developed by the Technical University of Denmark (DTU) can be found online (<http://www.cbs.dtu.dk/services/NetNES/>). It predicts aminoacids that could be involved in a Nuclear Exclusion Sequence (NES)

prolactine non sécrétée adopte au sein d'une cellule cancéreuse pulmonaire (**Figures 52 et 53**). Une analyse de cette lignée en Western blot montre que l'anticorps anti-GFP reconnaît deux bandes distinctes (**Figure 52A**). La bande supérieure a une taille apparente légèrement inférieure à 55 Da, correspondant à la taille attendue de la protéine de fusion PRL (25 kDa)-GFP (27 kDa). La bande inférieure, moins abondante, a une taille apparente de 43 kDa. Etant donné que la GFP est fusionnée à l'extrémité amino-terminale de la prolactine, nous avons conclu que la prolactine devait être tronquée de sa partie carboxy-terminale. Le clivage protéolytique de la prolactine en C-terminal a déjà été observé *in vivo* (**voir Contexte, paragraphe 1.3.1**). La taille de ce fragment tronqué de la prolactine doit être d'environ 16 kDa (43 kDa - 27 kDa).

Dans le même temps, j'ai observé que la lignée stable H1299 GFP-PRL perdait peu à peu sa fluorescence au fur et à mesure des passages (**Figure 52B**), un phénomène non observé dans la lignée contrôle H1299 GFP. Alors que la fluorescence due à la protéine de fusion GFP-PRL était observée dans 95% des cellules après sélection et tri cellulaire, cette proportion tombe à 15% après un mois de culture dans un milieu où la pression de sélection est maintenue. L'analyse du cycle cellulaire en cytométrie de flux montre également que les cellules ne prolifèrent plus autant, avec 7% de cellules en phase S contre 32% auparavant, pour un même degré de confluence. Nous nous sommes alors demandés si la perte de croissance cellulaire de la lignée H1299 GFP-PRL ne pourrait pas être due au clivage observé de la prolactine. En effet, il a été décrit que la prolactine 16 kDa peut jouer un rôle anti-prolifératif, sur les cellules endothéliales notamment, en augmentant la transcription de régulateurs négatifs du cycle cellulaire tels que *p21* ([Tabruyn et al., 2005](#)). Lorsque que l'on transfecte transitoirement des cellules H1299 avec une construction codant la forme courte GFP-PRL16, on observe bien la protéine correspondante en western blot (**Figure 52C, panneau de gauche**). Mais l'on se révèle incapable d'établir une lignée stable : même après sélection, une très faible proportion des cellules (environ 5%) expriment la GFP-PRL16. Néanmoins, ce rôle anti-prolifératif de la PRL16 ne semble pas être dû à l'activation transcriptionnelle de *p21* (**Figure 52C, panneau de droite**). En effet, on ne détecte pas d'augmentation d'expression du gène *p21* en qPCR lors de la transfection de GFP-PRL16, alors que son induction est bien observée suite à un traitement avec la TSA (contrôle positif). Etant donné que la forme clivée de la prolactine semble tout de même limiter la croissance cellulaire, nous nous sommes demandé comment les cellules de la lignée stable H1299 GFP-PRL réussissaient à outrepasser en partie ce rôle antiprolifératif. Alors que la GFP-PRL16 clivée est relativement abondante suite à une transfection transitoire, j'ai observé que la bande correspondante était moins intense après établissement de la lignée stable (**Figure 52D**). Cela pourrait suggérer que les cellules H1299 GFP-PRL ont réussi à inhiber en partie le clivage de la prolactine en une forme qui limite leur prolifération.

2.5.3. L'acétylation de la prolactine non sécrétable dans notre lignée modèle de cancer du poumon contrôle sa localisation subcellulaire

Nous avons décrit la rétention intracellulaire subie par la GFP-PRL non sécrétable (**Figure 50C**). J'ai de plus observé que la localisation subcellulaire de la prolactine varie en fonction de la lignée dans laquelle elle est transfectée (**Figure 53A**). Alors que les protéines de fusion GFP-PRL ou la HA-PRL (détectée en immunofluorescence avec un anticorps anti-HA) exprimées par les cellules H1299 sont localisées dans l'ensemble de la cellule (avec un signal nucléaire plus important que le signal cytoplasmique), ces protéines adoptent une localisation cytoplasmique uniquement dans les cellules COS. Parmi les nombreuses caractéristiques qui distinguent les cellules COS des cellules H1299, il en est une qui nous a particulièrement intéressés : les cellules COS sont connues pour exhiber un haut niveau d'acétylation des protéines. Par une analyse en western-blot, j'ai pu montrer que la prolactine elle-même est acétylée dans ces deux types cellulaires (**Figure 53B**). En effet, les deux formes (longue et clivée) sont détectées par un anticorps reconnaissant spécifiquement les lysines acétylées. J'ai ensuite soumis les cellules COS et H1299 à diverses conditions modulant le niveau global d'acétylation dans la cellule (**Figure 53C**). Les traitements par des inhibiteurs de déacétylases (TSA et sodium butyrate) et la surexpression de l'acétylase p300 conduisent à une

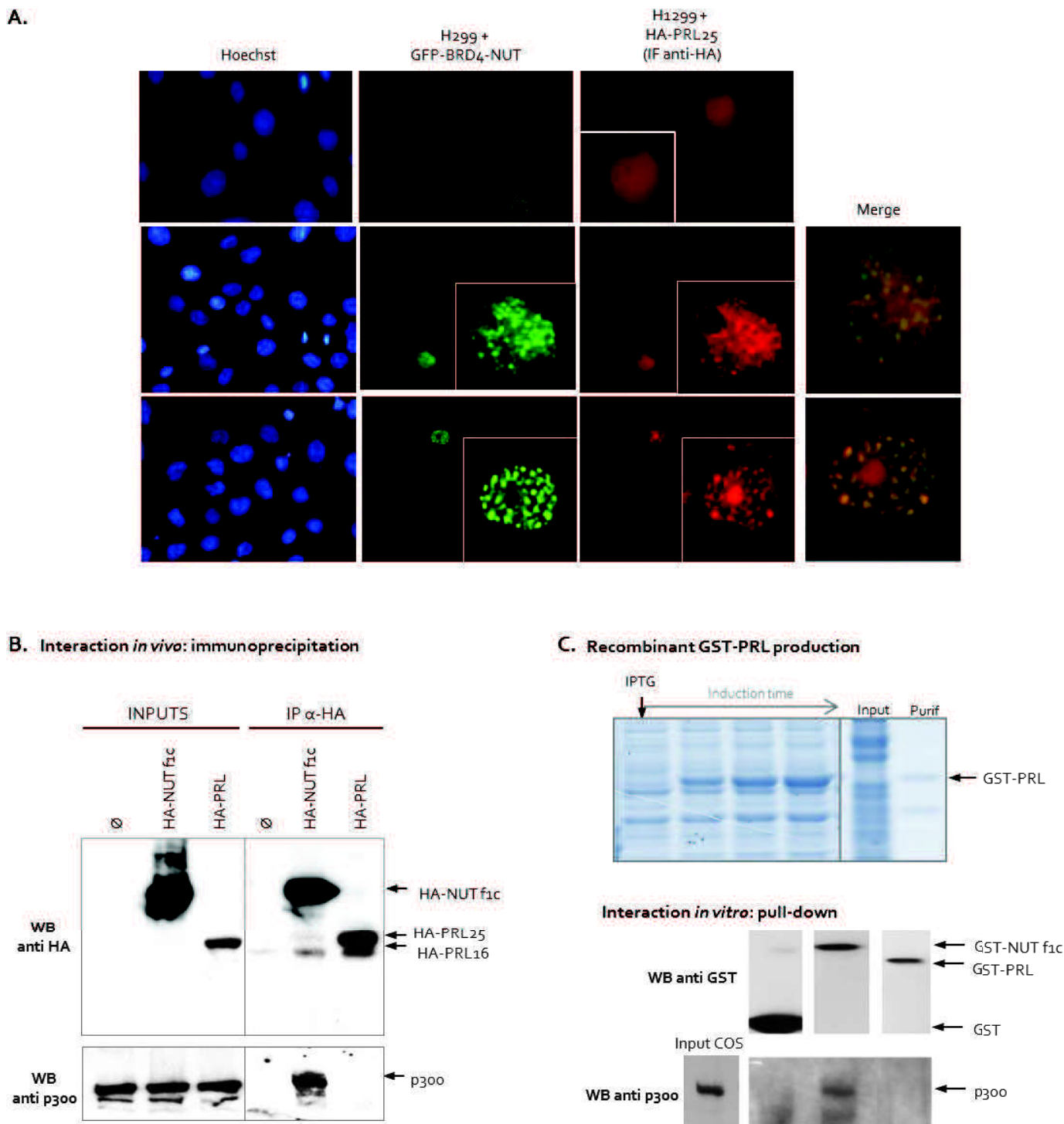


Figure 54: PRL does not physically interact with the histone acetyltransferase p300

A. Co-transfection of HA-PRL and GFP-BRD4-NUT fusion proteins in H1299 cells. BRD4-NUT protein is localized in nuclear foci. When only HA-PRL is expressed within a cell (top panel), HA-PRL protein (detected by immunofluorescence with an anti-HA antibody) localizes homogeneously in the nucleus. When HA-PRL and BRD4-NUT are expressed simultaneously within a cell (bottom panels), PRL protein adopts also a punctuated nuclear localization.

B. Physical *in vivo* interaction between PRL protein and the p300 HAT is tested by immunoprecipitation. HA-PRL and HA-NUT f1c (previously described p300 interactant used as a positive control) are immunoprecipitated with anti-HA antibody (top panel). The presence of endogenous p300 protein in the inputs and immunoprecipitated samples is detected with an anti-p300 antibody (bottom panel).

C. Physical *in vitro* interaction between PRL protein and the p300 HAT is tested by GST pull-down. Bacterial production of GST (negative control), GST-PRL and GST-NUT f1c (positive control) recombinant proteins is induced by IPTG (top panel). Recombinant proteins are purified on beads and can be detected with an anti-GST antibody (middle panel). Finally, beads are incubated with a lysate of COS cells and p300 endogenous protein in input and pull-down sample is detected with an anti-p300 antibody (bottom panel).

augmentation de l'acétylation, alors qu'un traitement par un inhibiteur de p300 conduit à sa diminution. J'ai alors observé une modification de la localisation subcellulaire de la prolactine en fonction du niveau d'acétylation. Quand celui-ci est élevé, la GFP-PRL tend à être exclue du noyau et à s'accumuler dans le cytoplasme. A l'inverse, on observe une relocalisation nucléaire de la GFP-PRL suivant un traitement qui diminue le niveau d'acétylation. Ceci suggère que l'acétylation de la prolactine puisse directement contrôler sa localisation au sein de la cellule cancéreuse. Une explication possible aux figures d'exclusion nucléaire de la prolactine observées lorsque le niveau d'acétylation est élevé est la présence au sein de la séquence protéique de la prolactine d'une séquence consensus pour un signal d'exclusion nucléaire (Nuclear Exclusion Sequence - NES) (**Figure 53D**). Nous pouvons proposer que l'acétylation de la prolactine puisse rendre ce signal accessible et/ou fonctionnel et conduire à sa relocalisation cytoplasmique.

2.5.4. La prolactine non-sécrétable pourrait être liée aux voies cellulaires d'acétylation, mais pas via une interaction directe avec l'acétyltransférase p300

Ces différents résultats ont mis en lumière le rôle potentiel de p300 sur la prolactine dans les cellules cancéreuses pulmonaires. p300 pourrait être l'enzyme responsable de l'acétylation de la prolactine, étant donné que la surexpression et l'inhibition spécifique de p300 s'avèrent capable de moduler sa localisation, mais aussi son clivage (données non présentées). Dans un premier temps, j'ai cherché à savoir si la prolactine et p300 étaient colocalisés (**Figure 54A**). Je me suis appuyée pour cela sur un contexte un peu particulier où la localisation de p300 est modifiée et restreinte à quelques foci nucléaires. Ceci est l'effet produit par la présence dans la cellule de la protéine de fusion BRD4-NUT exprimée suite à une translocation dans les carcinomes de la ligne médiane (NUT midline carcinoma – NMT) ([Reynoird et al., 2010b](#)). Le bromodomaine de la protéine BRD4 permet le ciblage de p300 à des histones acétylées via le domaine NUT de la protéine de fusion, causant la propagation de l'acétylation par la HAT p300 aux nucléosomes voisins. Cette boucle amplificatrice cause la rétention des enzymes p300 cellulaires dans des foci nucléaires hyperacétylés. Par une co-transfection de BRD4-NUT et de HA-PRL, j'ai voulu savoir si la prolactine allait se relocaliser dans les foci concentrant p300 (**Figure 54A**). C'est ce que nous observons, étant donné que la HA-PRL localisée de manière homogène dans le noyau lorsqu'elle est exprimée seule au sein d'une cellule (**panneau du haut**), présente une localisation ponctuelle lorsqu'exprimée conjointement à BRD4-NUT, dans des foci qui colocalisent avec les foci BRD4-NUT retenant p300 (**panneaux du bas**).

J'ai ensuite tenté de démontrer l'existence d'une interaction physique entre la prolactine et p300 dans les cellules cancéreuses (**Figure 54B et C**). Une telle interaction, si elle était avérée, pourrait fournir des bases mécanistiques à l'action de la prolactine. On sait en effet que des modifications de la localisation et/ou de l'activité de p300 peuvent conduire directement à la transformation cellulaire, comme c'est le cas dans les tumeurs BRD4-NUT où la rétention de p300 dans les foci nucléaires est directement oncogène ([Reynoird et al., 2010b](#)). J'ai d'abord tenté de détecter cette interaction *in vivo* par immunopurification (**Figure 54B**). La transfection de NUT-F1c, fragment f1c de la protéine NUT connu pour interagir avec p300, fournit un contrôle positif : en fixant spécifiquement HA-NUT-f1c avec un anticorps anti-HA, on est capable d'immunoprécipiter conjointement la protéine p300 endogène. Mais ceci n'est pas observable avec la prolactine : lorsque l'on fixe spécifiquement HA-PRL, on s'avère incapable de détecter p300 dans l'échantillon immunoprécipité. Je me suis alors tournée vers un test *in vitro* avec prolactine recombinante GST-PRL produite en système bactérien (**Figure 54C**). Lorsque la GST-PRL, fixée sur billes, est incubée avec un lysat cellulaire, elle s'avère incapable de lier la protéine p300 cellulaire, comme c'est le cas pour notre contrôle positif GST-NUT-f1c. Ainsi, je n'ai pas été en mesure de vérifier expérimentalement notre hypothèse d'une interaction physique entre p300 et PRL. Bien que ce modèle cellulaire H1299 GFP-PRL nous donne quelques pistes sur le comportement d'une prolactine non sécrétable dans les cellules cancéreuses pulmonaires (clivage, acétylation, changements de localisation), les mécanismes sur lesquels repose l'action de la prolactine dans les cellules cancéreuses CBPC restent inconnus.

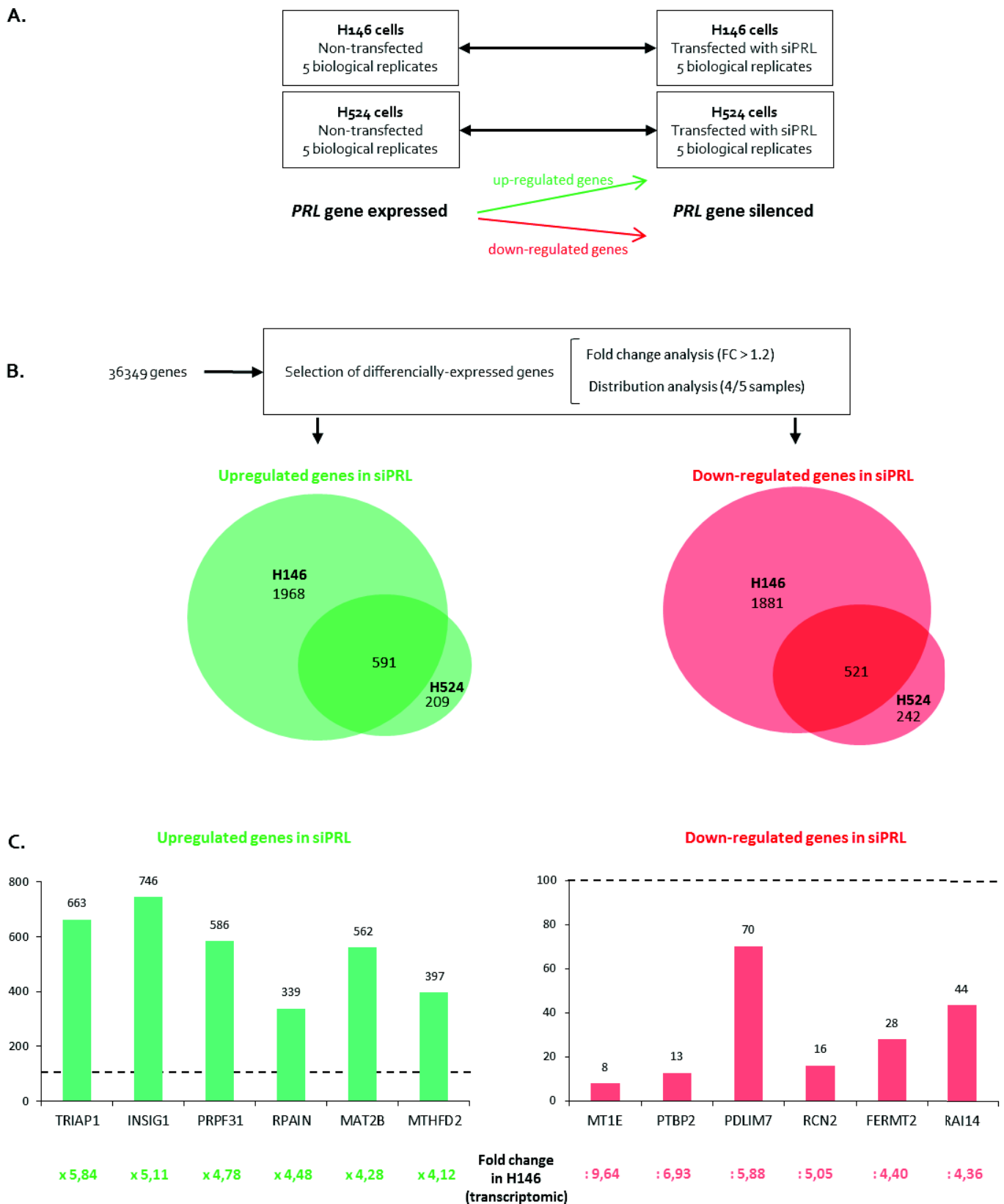


Figure 55: Differential transcriptome in *PRL*-silenced versus *PRL*-expressing SCLC cells

A. Principles of transcriptomic data analysis. In two different *PRL*-positive SCLC cell lines (H146 and H524), we compared the transcriptomes of the untreated cell line and the siPRL-transfected cell line, in order to identify up-regulated or down-regulated genes following *PRL* silencing.

B. After fold change and distribution analysis, we selected the probesets that were differentially expressed in 4 samples out of 5 biological replicates, with a fold change greater than 1.2. Finally, we transposed probesets results in term of genes results. The numbers of differentially expressed genes in both cell lines are repressed by Venn diagrams.

C. Validation of the 12 differentially-expressed genes (common to both cell lines) that exhibit the greatest fold-changes. Expression in siPRL condition is obtained by qPCR, in one biological replicate of H146 cells, by normalization with *Actin* and *U6*. 100% = Expression ratio in non-transfected *PRL*-expressing H146 cells. Fold change observed in transcriptomic data is indicated.

2.6. Les analyses haut débit des signatures géniques associées à l'expression du gène *PRL* pourraient aider à décrypter son mode d'action

Pour tenter de comprendre les mécanismes cellulaires liés à l'expression ectopique du gène *PRL* et son implication dans l'acquisition d'une résistance accrue par les cellules cancéreuses, nous avons adopté deux méthodes d'analyse à haut débit de l'expression des gènes : l'une portant sur les cellules de CBPC exprimant ou non le gène *PRL*, et l'autre portant sur les tumeurs de patients *PRL*+ et *PRL*-.

2.6.1. L'expression de *PRL* régule un large groupe de gènes dans les cellules de cancer du poumon à petites cellules

Pour identifier les gènes dont l'expression est affectée par l'activation ectopique du gène *PRL* dans les cellules cancéreuses pulmonaires, j'ai lancé une étude à grande échelle des expressions géniques dans les lignées de cancer du poumon à petites cellules (**Figures 55 à 57**). Une analyse transcriptomique (technologie Illumina) a été réalisée par la plate-forme du CHU d'Angers dirigée par le Dr Guardiola (**voir Méthodes**). Elle portait sur les ARN extraits des lignées *PRL*-positives H146 et H524, ainsi que sur les ARN de ces mêmes lignées dans lesquelles l'expression de *PRL* a été diminuée par transfection conjointe des deux siARN *PRL* (**Figure 55A**). Nous avons ensuite identifié les gènes différentiellement exprimés entre ces deux conditions : les gènes up-régulés dans la condition siPRL (correspondant à des gènes réprimés, directement ou indirectement, lorsque le gène *PRL* s'exprime) d'une part, et les gènes down-régulés dans la condition siPRL (correspondant à des gènes activés lorsque le gène *PRL* s'exprime) d'autre part (**Figure 55B**). Cette approche nous a permis d'identifier 5431 gènes différentiellement exprimés, dont 1112 sont communs aux deux lignées cellulaires étudiées. Parmi ceux-ci, un sous-groupe de 591 gènes est surexprimé suite au silençage de *PRL*. A l'inverse, 521 gènes voient leur expression diminuer lorsque *PRL* est ciblé par le siARN (**Figure 55B**). Ces résultats ont été confirmés en qPCR pour les douze gènes présentant les plus grandes variations d'expression (**Figure 55C**).

Nous avons ensuite tenté de déterminer les grandes fonctions cellulaires associées aux gènes identifiés comme différentiellement exprimés en réalisant une analyse d'enrichissement des termes de Gene Ontology (**Figure 56**). Certains compartiments cellulaires et fonctions biologiques apparaissent sur-représentés parmi les gènes différentiellement exprimés. Dans les gènes up-régulés suite au siPRL, on observe un enrichissement des termes liés au compartiment nucléaire. Dans les gènes down-régulés suite au siPRL, on observe un enrichissement des termes concernant les chromosomes mitotiques. On peut noter qu'aucun terme lié à la membrane plasmique ou au compartiment extracellulaire n'est significativement enrichi. Lorsque l'on s'intéresse aux fonctions cellulaires et processus biologiques assurés par les gènes différentiellement exprimés, on retrouve de grandes catégories particulièrement enrichies (**Figure 56**) : les activités liées à l'expression du génome ou au métabolisme pour les gènes surexprimés lors du silençage de *PRL*, ou le catabolisme des protéines et le cycle cellulaire pour les gènes dont l'expression diminue sous siPRL.

Pour préciser davantage le rôle que pourrait jouer l'expression de *PRL* au sein des cellules cancéreuses pulmonaires, nous avons réalisé une analyse GSEA (Gene Set Enrichment Analysis) (**Figure 57**) permettant de relier les gènes différentiellement exprimés que nous avons identifiés à des groupes de gènes issus d'études transcriptomiques publiques différentiellement exprimés entre deux états biologiques. Les gènes que nous avons identifiés comme up-régulés après siPRL (**panneau de gauche**), c'est-à-dire les gènes qui étaient moins exprimés lorsque *PRL* était exprimé et qui pourraient potentiellement être directement ou indirectement inhibés par l'expression de *PRL*, sont enrichis dans des groupes de gènes impliqués dans la maturation des ARN, la réparation de l'ADN ou le métabolisme. On identifie aussi des sous-groupes de gènes qui sont down-régulés dans les tumeurs agressives, ainsi que des gènes dont l'expression est réduite suite à l'application d'un stress sur les cellules tumorales (sous la forme d'une hypoxie ou de traitement anti-cancéreux variés).



Cellular component			Upregulated genes in siPRL	Down-regulated genes in siPRL
Cellular component	Nucleus	0005634 nucleus	215 / 157	173 / 140
		0044428 nuclear part	105 / 58	62 / 51
		0031981 nuclear lumen	84 / 48	41 / 41
		0005730 nucleolus	47 / 22	24 / 20
	Mitotic chromosomes	0000793 condensed chromosome	1 / 4	12 / 4
		0000939 inner kinetochore of condensed chromosome	0 / 0	2 / 0
	Organelles	0044446 intracellular organelle part	189 / 130	145 / 115
		0043231 intracellular membrane-bounded organelle	338 / 246	268 / 219
		0044446 intracellular organelle part	189 / 130	145 / 115
		0070013 intracellular organelle lumen	103 / 58	50 / 49
		0005739 mitochondrion	55 / 34	37 / 30
	Cytoplasm	0005737 cytoplasm	288 / 228	284 / 203
		0044444 cytoplasmic part	184 / 152	182 / 135
Biological process	Genome expression	0010467 gene expression	140 / 95	74 / 84
		0034470 ncRNA processing	20 / 6	4 / 5
		0006396 RNA processing	39 / 18	13 / 16
		0003723 RNA binding	45 / 23	22 / 21
	Metabolism	0044237 cellular metabolic process	265 / 211	197 / 187
		0044238 primary metabolic process	268 / 219	202 / 195
		0016070 RNA metabolic process	59 / 31	24 / 27
		0034641 cellular nitrogen compound metabolic process	161 / 117	108 / 104
		0034660 ncRNA metabolic process	23 / 8	5 / 7
		0090304 nucleic acid metabolic process	134 / 98	89 / 87
	Cell cycle	0000278 mitotic cell cycle	14 / 12	29 / 11
		0007049 cell cycle	25 / 25	45 / 23
		0048285 organelle fission	9 / 8	20 / 7
		0031577 spindle checkpoint	0 / 0	5 / 0
		0007067 mitosis	7 / 7	18 / 7
		0022403 cell cycle phase	15 / 14	27 / 12
		0000070 mitotic sister chromatid segregation	1 / 1	7 / 1
		0007094 mitotic cell cycle spindle assembly checkpoint	0 / 0	4 / 0
		0022402 cell cycle process	19 / 19	32 / 17
		0030071 regulation of mitotic metaphase/anaphase transition	0 / 1	5 / 1
	Protein catabolism	0030163 protein catabolic process	26 / 20	33 / 18
		0019941 modification-dependent protein catabolic process	20 / 18	32 / 16
		0070647 protein modification by small protein conjugation or removal	7 / 5	14 / 5
		0016567 protein ubiquitination	5 / 4	12 / 3
	Proteins and organelles intracellular localization	0015031 protein transport	48 / 25	31 / 22
		0008104 protein localization	53 / 29	39 / 26
		0046907 intracellular transport	42 / 22	35 / 20
		0033036 macromolecule localization	57 / 36	48 / 32
		0051641 cellular localization	47 / 31	54 / 28
		0051649 establishment of localization in cell	44 / 29	47 / 26
		0051640 organelle localization	5 / 3	11 / 3
		0051656 establishment of organelle localization	4 / 2	9 / 2
		0006996 organelle organization	53 / 44	61 / 39

Figure 56: Cellular components and biological processes affected by the silencing of *PRL* are identified by a Gene Ontology analysis

An analysis of Gene Ontology is performed with the AMEN software (Chalmel and Primig, 2008).

The analysis compares the theoretical number of genes expected for each term of gene ontology, and the observed number of genes affected by the siPRL. We present the significantly enriched terms.

Les gènes que nous avons identifiés comme down-régulés après siPRL (**panneau de droite**), c'est-à-dire les gènes qui étaient davantage exprimés lorsque *PRL* était exprimé et qui pourraient potentiellement être directement ou indirectement activés par l'expression de *PRL*, sont enrichis dans des groupes de gènes up-régulés dans les tumeurs et dans les tumeurs agressives en particulier, mais aussi dans des groupes de gènes liés à l'adhésion et la migration cellulaire ou up-régulés dans les tumeurs invasives. On identifie aussi des sous-groupes de gènes dont l'expression est augmentée suite à l'application d'un stress sur les cellules tumorales (sous la forme d'une hypoxie ou de traitement anti-cancéreux). Enfin, on identifie de nombreuses occurrences de gènes impliqués dans l'acétylation.

2.6.2. Des cibles potentielles de l'expression de *PRL* sont identifiées dans les transcriptomes de patients

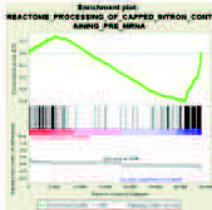
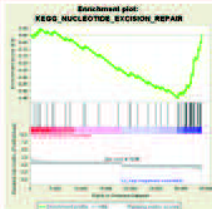
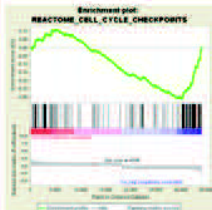
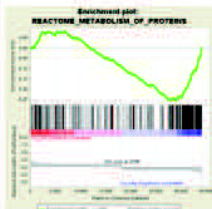
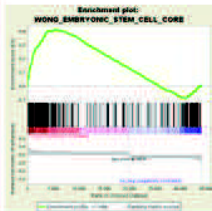
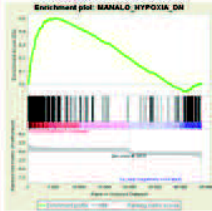

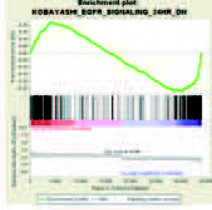
Grâce aux données transcriptomiques de 253 patients de la cohorte (obtenues par la technologie Affymetrix), nous avons réalisé une analyse supervisée des transcriptomes des patients *PRL*-positifs (exhibant une expression aberrante du gène *PRL* au sein de leur tumeur) et des patients *PRL*-négatifs (**Figure 58**). Ce travail nous a permis d'identifier 2457 gènes différentiellement exprimés entre ces deux types de tumeurs : 1348 gènes up-régulés dans les tumeurs *PRL*-positives (correspondant à des gènes potentiellement activés, directement ou indirectement, lorsque le gène *PRL* s'exprime) et 1109 gènes down-régulés dans ces mêmes tumeurs *PRL*-positives (correspondant à des gènes potentiellement réprimés, de manière directe ou indirecte, par l'expression de *PRL*) (**Figure 58A**).

Nous avons ensuite procédé à l'analyse des termes de Gene Ontology associés à ces deux sous-groupes de gènes différentiellement exprimés (**Figure 58B**). Parmi les gènes surexprimés dans les tumeurs *PRL*+ (**panneau de gauche**), nous retrouvons quasi-exclusivement des termes ayant trait au cycle cellulaire, à la mitose, à la réplication de l'ADN, à la ségrégation des chromosomes et au fuseau de division. Aux gènes exprimés plus faiblement dans les tumeurs *PRL*+ que dans les tumeurs *PRL*- (**panneau de droite**) on associe une abondance de termes liés à la membrane plasmique, aux stimuli extracellulaires et à la transduction des signaux, aux mécanismes de défense et à la réponse immune.

Une analyse de corégulation nous a permis d'autre part d'identifier 56 gènes corégulés avec *PRL* chez les patients (**Figure 59A**). Ces gènes présentent un niveau d'expression qui est positivement ou négativement corrélé avec le niveau d'expression du gène *PRL* dans les 27 échantillons de tumeurs, ce qui pourrait signer une action directe de *PRL* sur ces gènes.

Enfin, en analysant la heat map des gènes différentiellement exprimés dans les tumeurs *PRL*+ versus *PRL*- (**Figure 58A**), nous avons observé l'existence d'un sous-groupe de patients considérés comme *PRL*- négatifs, mais dont la signature génique est très proche de celle des tumeurs *PRL*+ (**Figure 59B**). Ces patients, que nous avons renommés « *PRL*-/*PRL*+ like », possèdent un pronostic plutôt sombre, significativement différent du pronostic des patients *PRL*- *sensu stricto*, et proche du pronostic des patients *PRL*+. Nous proposons que ces patients *PRL*-/*PRL*+ like puissent correspondre à des patients considérés de manière erronée comme *PRL*-négatifs par les analyses transcriptomiques, mais qui sont en réalité des patients *PRL*-positifs. C'est notamment le cas du patient n°61, qui avait été considéré comme *PRL*- en transcriptomique (**Figure 35**), qui s'était révélé *PRL*+ en PCR (**Figure 49**) car exprimant une forme sévèrement tronquée de l'ARNm *PRL*, et qui s'avère faire partie de ce sous-groupe de patients *PRL*-/*PRL*+ like du fait d'une signature génique proche de celles des patients *PRL*+

Genes upregulated in siPRL (Genes repressed when *PRL* is expressed)

Pathway	Example of enriched genes subset	Other enriched genes subset
Genes involved in RNA transcripts maturation (splicing, capping, transport)	 <p>Genes involved in Processing of Capped Intron-Containing Pre-mRNA</p>	REACTOME_FORMATION_AND_MATURATION_OF_MRNA_TRANSCRIPT KEGG_SPLICEOSOME REACTOME_ELONGATION_AND_PROCESSING_OF_CAPPED_TRANSCRIPTS REACTOME_MRNA_SPLICING REACTOME_METABOLISM_OF_RNA REACTOME_TRANSPORT_OF_FROM_AN_INTRON_CONTAINING_TRANSCRIPT REACTOME_METABOLISM_OF_MRNA REACTOME_MRNA_PROCESSING
Genes involved in DNA repair	 <p>Genes involved in Nucleotide excision repair (NER): repair of bulky DNA damage caused by environmental carcinogens, and exposure to UV-light</p>	REACTOME_DNA_REPAIR REACTOME_TRANSCRIPTION_COUPLED_NER KEGG_NUCLEOTIDE_EXCISION_REPAIR REACTOME_GLOBAL_GENOMIC_NER KEGG_BASE_EXCISION_REPAIR KEGG_MISMATCH_REPAIR REACTOME_P53_INDEPENDENT_DNA_DAMAGE_RESPONSE
Genes involved in cell cycle ckeckpoints	 <p>Genes involved in Cell Cycle Checkpoints</p>	REACTOME_G2_M_CHECKPOINTS
Genes involved in protein metabolism	 <p>Genes involved in Metabolism of proteins</p>	KEGG_OXIDATIVE_PHOSPHORYLATION KEGG_PROTEIN_EXPORT
Genes of ES cells	 <p>The 'core ESC-like gene module': genes up-regulated in embryonic stem cells (ESC)</p>	BENPORATH_ES_CELLS
Genes down-regulated under hypoxia and deprivations	 <p>Genes down-regulated in response to both hypoxia and overexpression of an active form of HIF1A [GenelD=3091].</p>	ELVIDGE_HYPOXIA_DN ELVIDGE_HIF1_KNOCKDOWN_TARGETS_UP PENG_LEUCINE_DEPRIVATION_DN CROONQUIST_IL6_DEPRIVATION_DN
Genes down-regulated in tumors, aggressive tumors or metastatic tumors	 <p>Genes down-regulated in breast cancer tumors having type 1 amplification in the 20q13 region associated with a poor prognosis.</p>	LI_WILMS_TUMOR_VS_FETAL_KIDNEY_1_DN WELCSH_BRCA1_DOWN
Genes down-regulated after anti-tumor treatments	 <p>Genes down-regulated in H1975 cells (non-small cell lung cancer, NSCLC) after treatment with EGFR inhibitor CL-387785 [PubChem=2776] for 24h.</p>	ZHANG_RESPONSE_TO_CANTHARIDIN_DN PENG_RAPAMYCIN_RESPONSE_DOWN FRASOR_RESPONSE_TO_SERM_TAMOXIFEN_DN

Down in siPRL (Genes expressed when PRL is expressed)

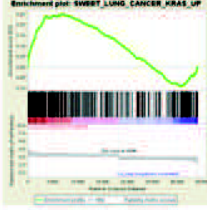


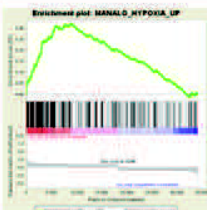
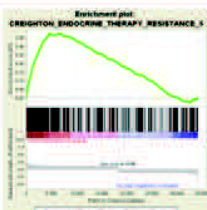
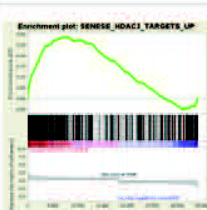
Pathway	Example of enriched genes subset	Other enriched genes subset
Genes up-regulated in tumors	 <p>Genes up-regulated in the Kras2LA mouse lung cancer model with mutated KRAS [GeneID=3845].</p>	DELYS_THYROID_CANCER_UP GRUETZMANN_PANCREATIC_CANCER_UP HOSHIDA_LIVER_CANCER_SUBCLASS_S1 CAIRO_HEPATOBLASTOMA_UP ZHANG_BREAST_CANCER_PROGENITORS_UP WELCSH_BRCA1_TARGETS_1_UP GRADE_COLON_AND_RECTAL_CANCER_UP WAMUNYOKOLI_OVARIAN_CANCER_GRADES_1_2_UP VECCHI_GASTRIC_CANCER_EARLY_UP WANG_LMO4_TARGETS_UP LI_WILMS_TUMOR_VS_FETAL_KIDNEY_1_UP
Genes up-regulated in aggressive tumors	 <p>Cluster 6 of method A: up-regulation of these genes in patients with non-small cell lung cancer (NSCLC) predicts poor survival outcome.</p>	LEE_LIVER_CANCER_POOR_SURVIVAL_UP CHIANG_LIVER_CANCER_SUBCLASS_PROLIFERATION_UP WOO_LIVER_CANCER_RECURRENCE_UP BORCZUK_MALIGNANT_MESOTHELIOMA_UP KOYAMA_SEMA3B_TUMOR_SUPPRESSOR_TARGETS_DN RICKMAN_TUMOR_DIFFERENTIATED_WELL_VS_POORLY_DN BERTUCCI_MEDULLARY_VS_DUCTAL_BREAST_CANCER_DN LANDIS_ERBB2_BREAST_TUMORS_324_UP
Genes involved in cell migration or up-regulated in metastatic tumors	 <p>Genes up-regulated in endometroid endometrial tumors from patients with lymph node metastases compared to those without the metastases.</p>	WU_CELL_MIGRATION KEGG_REGULATION_OF_ACTIN_CYTOSKELETON KEGG_TIGHT_JUNCTION BIDUS_METASTASIS_UP ALONSO_METASTASIS_UP PROVENZANI_PRIMARY_VS_METASTASIS_DN ONDER_CADHERIN_E_KNOCKDOWN_UP
Genes up-regulated under hypoxia	 <p>Genes up-regulated in response to both hypoxia and overexpression of an active form of HIF1A [GeneID=3091].</p>	ELVIDGE_HYPOXIA_UP JIANG_HYPOXIA_NORMAL_UP ELVIDGE_HIF1A_AND_HIF2A_KNOCKDOWN_DN ELVIDGE_HYPOXIA_BY_DMOG_UP
Genes up-regulated after anti-tumor treatment	 <p>The 'group 5 set' of genes associated with acquired endocrine therapy resistance in breast tumors</p>	TOOKER_RESPONSE_TO_BEXAROTENE_UP HELLER_HDAC_TARGETS_UP_RESPONSE_TSA
Genes linked to acetylation	 <p>Genes up-regulated in U2OS cells (osteosarcoma) upon knockdown of HDAC3 [GeneID=8841] by RNAi.</p>	HELLER_HDAC_TARGETS_UP_RESPONSE_TSA HELLER_HDAC_TARGETS_UP HELLER_HDAC_TARGETS_SILENCED_BY_METHYLATION_UP SENESE_HDAC3_TARGETS_UP SENESE_HDAC1_AND_HDAC2_TARGETS_UP

Figure 57: GSEA analysis of differential transcriptomes in SCLC PRL + versus PRL – identifies significantly concordant gene sets

A GSEA analysis was performed on differentially expressed genes identified by our transcriptomic analysis of H146 SCLC cells expressing *PRL* and silenced for *PRL* expression (siPRL). Some of the gene sets concordant with up- and down-regulated genes (enrichment in phenotype, category C2v3) are presented.

(Left panel) Genes up-regulated following siPRL

(Right panel) Genes down-regulated following siPRL

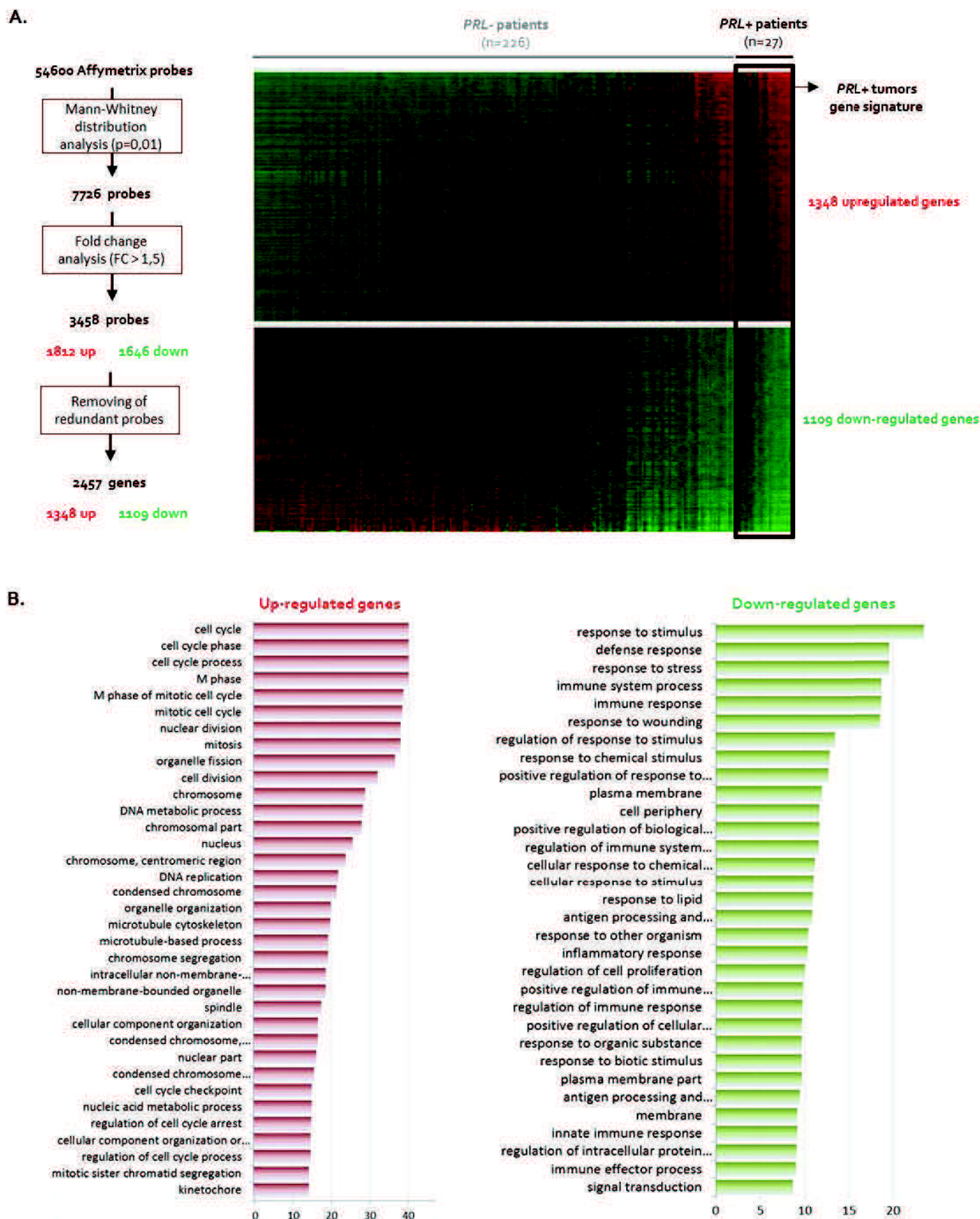


Figure 58: Differential transcriptome in *PRL*-expressing versus *PRL*-negative lung tumors

A. Principles of transcriptomic data analysis. Thanks to transcriptomic data from 253 patients of the cohort, we compared the transcriptomes of the *PRL*-positive and the *PRL*-negative lung tumors, in order to identify up-regulated or down-regulated genes following *PRL* ectopic expression. After a Mann-Whitney distribution analysis and a fold change analysis, we selected the probesets that were differentially expressed between samples, with a fold change greater than 1.5. Finally, we transposed probesets results in term of genes results. Differentially-expressed genes are represented by a heat map.

B. An analysis of Gene Ontology is performed in order to identify significantly enriched terms in up- and down-regulated subsets of genes. The $-\log$ corrected p -value is represented as an indicator of enrichment.

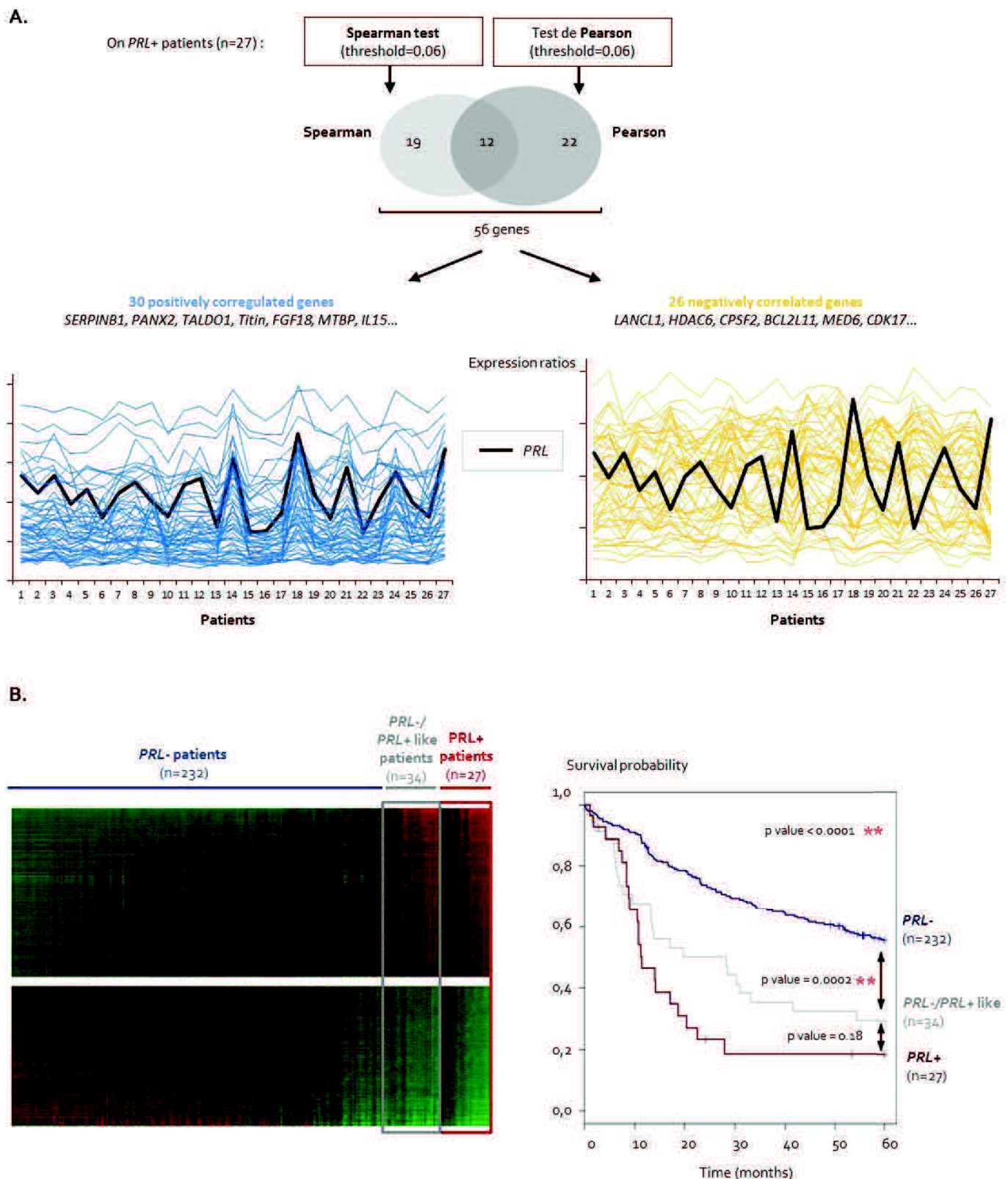
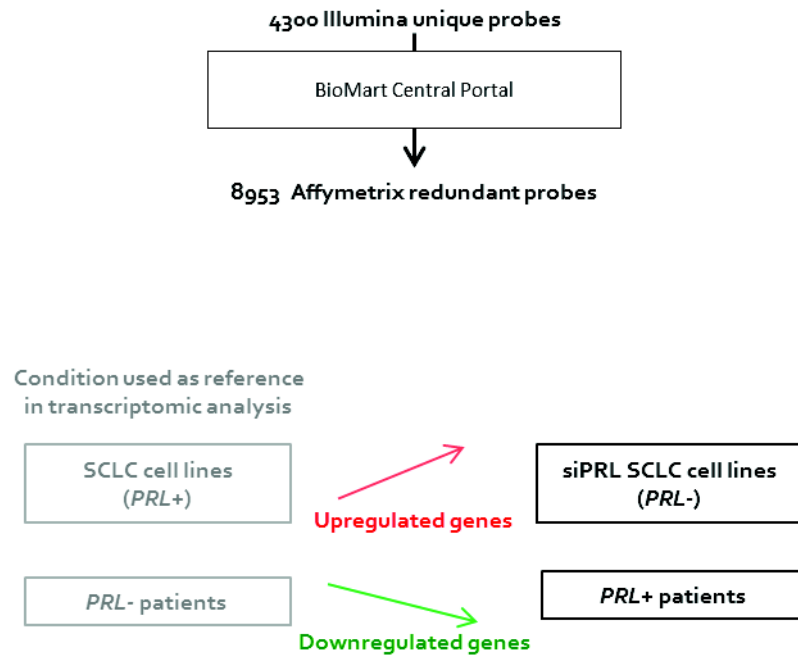


Figure 59: In-depth analysis of patients transcriptome: identification of *PRL*-coregulated genes and of *PRL*-/*PRL*+ like patients

A. Coregulation analysis was performed in order to identify genes whose expression levels are positively- (top panel) or negatively- (bottom panel) correlated with *PRL* expression levels with *PRL*-positive tumors. Two different coregulation approaches were performed with the Genespring software: the measure of similarity with Spearmann and Pearson tests.

B. 34 *PRL*- patients are identified because they have a gene signature closed to the *PRL*+ tumor signature (left panel) and are renamed *PRL*-/*PRL*+ like patients. Survival analysis of *PRL*-, *PRL*-/*PRL*+ like and *PRL*+ patients (right panel). The significance of the difference in prognosis is assessed by the univariate logrank test.

A.



B.

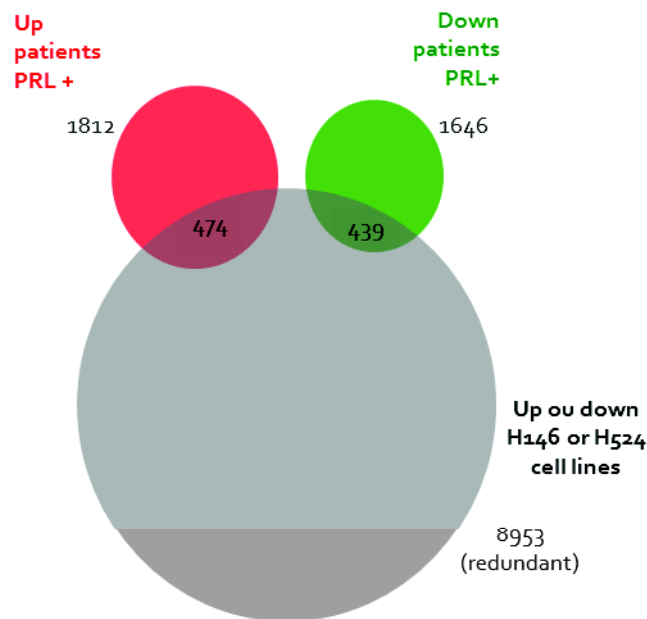


Figure 60: Comparison of patients and cell lines differentially expressed-genes

A. The transcriptomic analysis performed on cell lines and patients have not been made with the same technology (respectively Illumina and Affymetrix technologies). To be able to compared differentially-expressed genes in cell lines and patients, we need to match the two types of probes used. Thanks to the Biomart Central Portal available online (<http://central.biomart.org/>), we converted the 4300 non-redundant Illumina probes corresponding to the differentially-expressed genes in H146 and H524 cell lines (Figure 55) in 8953 redundant Affimatrix probes.

B. Comparison of redundant probes differentially expressed in patients and cell lines identifies a core of *PRL*-dependent genes.

Finalement, nous avons souhaité combiner les deux approches haut débit, celle sur lignées cellulaires et celle sur patients, pour tenter d'identifier un cœur de gènes fermement liés à l'expression de *PRL* (**Figure 6o**). Ce travail encore en cours a nécessité d'établir une correspondance entre sondes Illumina et sondes Affymetrix, étant donné que les analyses transcriptomiques sur lignées cellulaires et sur patients n'ont pas été réalisées sur la même plate-forme technologique (**Figure 6oA**). D'autre part, il faut prendre en compte le fait que les conditions prises comme références dans nos analyses (cellules non traitées par le siARN *PRL* d'une part, patients *PRL*-négatifs d'autre part) correspondent pour la première au cas où le gène *PRL* s'exprime et pour la seconde au cas où le gène *PRL* ne s'exprime pas. Il faudra donc tester le recouvrement des gènes up-régulés dans les lignées avec les gènes down-régulés dans les patients, et inversement (**Figure 6oA**). La comparaison des gènes down-régulés dans le siPRL sur cellules CBPC et up-régulés dans les tumeurs *PRL*+ d'une part, et up-régulés dans le siPRL sur cellules CBPC et down-régulés dans les tumeurs *PRL*+ d'autre part (**Figure 6oB**), réalisée de manière préliminaire sur sondes redondantes, met en lumière un nombre de gènes communs aux deux approches relativement faible. Pour autant, ces gènes se révèlent être de très bons candidats pour une action directe de *PRL* et pourront être caractérisés plus en avant. Concernant la relativement faible proportion de gènes communs aux deux approches, nous pensons qu'elle a une base biologique, les deux modèles utilisés pour les études haut débit étant différents : on s'appuie en effet d'une part sur des lignées cellulaires qui exprimaient le gène *PRL*, dans lesquelles on supprime brusquement l'expression de ce gène et pour lesquelles on observe un effet à très court terme (2 jours de siARN), et d'autre part des tumeurs, ayant amorcé leur progression il y a des dizaines d'années, qui ont toujours exprimé le gène *PRL* et qui se sont probablement accoutumées à cette expression ectopique. Enfin, même si la correspondance en terme de gènes est limitée, nous avons mis en lumière une bonne concordance des fonctions biologiques. Ainsi, les processus liés au métabolisme par exemple sont enrichis dans les conditions où la prolactine ne s'exprime pas (lignées CBPC siPRL et patients *PRL*-) alors que l'on retrouve les fonctions biologiques liées au cycle cellulaire enrichies dans les conditions où la prolactine s'exprime (lignées CBPC si- et patients *PRL*+).

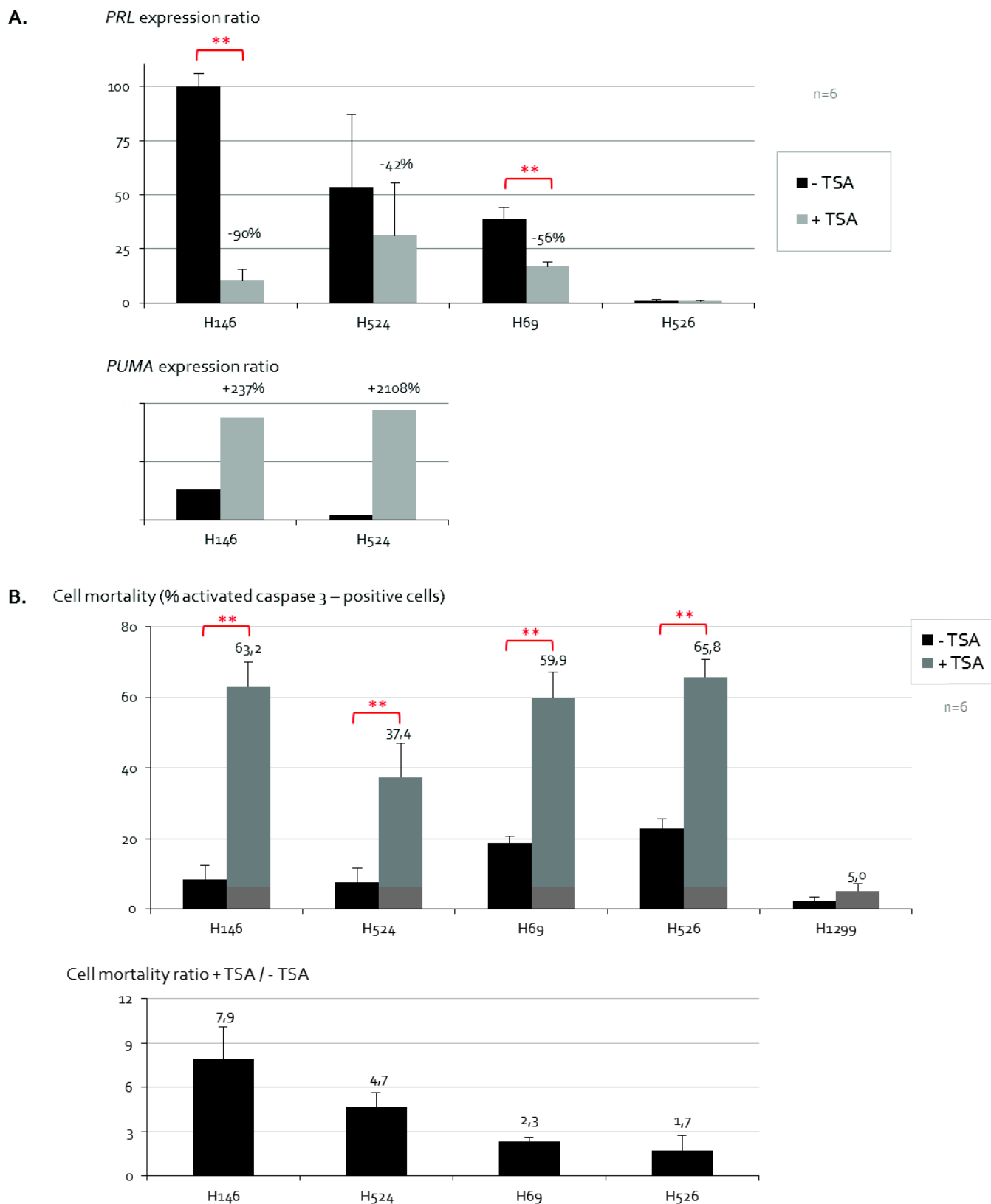


Figure 61: *PRL* silencing and *PRL*+ SCLC cell lines mortality induced by the HDAC inhibitor Trichostatin A (TSA)
SCLC cell lines were treated with 100 ng/mL TSA for 6 hours.

A. (Top panel) Expression ratios of *PRL* gene obtained from 6 independent RT-qPCR experiments. Expressions are normalized thanks to *Actin* control gene. 100% = *PRL* expression ratio in H146 si -. The significance of the difference in *PRL* expression between treated and untreated conditions is assessed by a Student t-test of mean comparison.

(Bottom panel) *PUMA* is used as a control gene for TSA treatment efficiency (*PUMA* is a known TSA-induced gene).

B. (Top panel) Cell mortality was measured by flow cytometry after incubation with an anti-activated Caspase 3 antibody in 6 independent experiments. (Bottom panel) Cell mortality ratio (Acti D-induced cell death versus basal cell death in control condition) was calculated for each of the 6 experiments. The significance of the difference in cell mortality between treated and untreated conditions is assessed by a Student t-test of mean comparison. * p value < 10% ** p value < 5%

2.7. L'expression aberrante du gène *PRL* peut être ciblée par des drogues épigénétiques

Nous avons montré que les tumeurs exprimant de manière ectopique le gène *PRL* sont des tumeurs agressives, causant des décès précoces (**Figure 38**). Nous avons également démontré que l'expression de *PRL* est directement associée à une chimiorésistance accrue, et que le silençage de l'expression de *PRL* permet de restaurer la sensibilité des cellules à la drogue et d'augmenter la mortalité induite par un stress génotoxique (**Figure 40**). Nous avons alors voulu proposer une thérapie potentielle ciblant les tumeurs *PRL*-positives, basée sur un traitement permettant de réduire le niveau d'expression du gène *PRL* dans les cellules cancéreuses pulmonaires.

Dans cette étude, nous avons également suggéré l'existence de liens entre l'expression de *PRL* et l'acétylation : une prolactine exogène est acétylée dans une lignée cancéreuse pulmonaire (**Figure 53B**), sa localisation intracellulaire est contrôlée par le niveau d'acétylation (**Figure 53C**), et des ensembles de gènes liés à l'acétylation ou aux drogues modulant l'acétylation apparaissent enrichis dans les analyses GSEA des transcriptomes de lignées transfectées par un siPRL (**Figure 57**).

Nous avons alors soumis les cellules de CBPC *PRL*-positives à un traitement par la trichostatine A (TSA), un inhibiteur des déacétylases (HDAC) ([Yoshida et al., 1990](#)). Un traitement à la TSA réduit de manière significative l'expression du gène *PRL* dans les lignées de cancer du poumon à petites cellules H146 et H69 (**Figure 61A**) : le niveau d'expression de *PRL* dans la lignée H146 est ainsi divisé par 10 après un traitement par la TSA. De plus, j'ai observé que les cellules *PRL*+ traitées à la TSA exhibent une importante mortalité, dépassant les 60% pour la lignée H146 (**Figure 61B, panneau du haut**). On remarque de plus que la mortalité TSA-induite (mesurée par le ratio mortalité sous TSA / mortalité basale en condition contrôle) a un profil tout à fait similaire au profil d'expression de *PRL* dans ces lignées : chez la lignée H146 qui exprime le plus le gène *PRL*, la mortalité est multipliée par 8 après traitement, alors qu'elle n'est que doublée dans la lignée H69 à expression plus faible (**Figure 61B, panneau du bas**). Ainsi, la sensibilité des cellules à la TSA apparait corrélée au niveau d'expression de *PRL* et un traitement par cet inhibiteur de HDAC pourrait permettre d'induire la mort cellulaire sélective des cellules tumorales exprimant le gène *PRL* de manière aberrante.

3. Discussion

3.1. La nécessité de nouvelles approches thérapeutiques dans la lutte contre les cancers pulmonaires

Jusqu'au milieu du siècle dernier, les maladies infectieuses (diphtérie, tétanos, tuberculose) étaient la principale cause des décès en France (Vallin and Meslé, 1988). Leur recul, dû à la généralisation des antibiotiques et des vaccinations, a entraîné une augmentation sensible de l'espérance de vie (Cambois et al., 2009). De ce fait, les efforts en termes de santé publique se sont naturellement déplacés vers de nouveaux terrains épidémiologiques. Diverses pathologies telles que les maladies neuro-dégénératives, le diabète et l'obésité, les maladies cardio-vasculaires ou bien encore les cancers ont pris une ampleur nouvelle dans les pays industrialisés du fait de l'allongement de la durée de vie, des progrès du diagnostic ou encore des modifications de notre environnement et de nos modes de vie.

Le cancer est une pathologie à laquelle chacun sera confronté, directement ou indirectement. On estime en effet qu'un homme a près d'une chance sur deux, et une femme plus d'une chance sur trois, de contracter un cancer au cours de sa vie (Menegoz and Chérié-Challine, 1998) et chaque Français sera à un moment donné affecté par la maladie d'un conjoint, d'un parent, d'un collègue ou d'un ami. Parmi les différentes pathologies tumorales, le cancer du poumon allie une incidence importante et un pronostic particulièrement sombre. Alors qu'il n'est pas le plus fréquent, c'est bien le plus mortel : il représente la première cause de mortalité par cancer en France et dans le monde. En conséquence, la lutte contre les cancers en général, et contre les cancers pulmonaires en particulier, est devenue un enjeu majeur de santé publique, donnant lieu à des actions internationales ou gouvernementales de grande ampleur (Projet GLOBOCAN de l'Organisation Mondiale de la Santé, Journée mondiale contre le cancer, les Plans Cancer en France depuis 2003), supportées par de nombreuses agences (Centre International de Recherche sur le Cancer, Union internationale contre le cancer, Institut National du Cancer) et relayées par d'innombrables associations (Association Internationale de Recherche sur le Cancer, La Ligue, l'ARC). Cette lutte s'appuie sur plusieurs leviers : études épidémiologiques, mise en place d'actions de prévention, développement du dépistage, prise en charge des malades et de leur famille – et bien sûr un axe de recherche.

La lutte contre les cancers pulmonaires s'appuie en premier lieu sur des mesures de prévention. La majeure partie des cancérogènes industriels et domestiques ayant été identifiés et interdits, les méthodes de prévention du cancer du poumon passent avant tout par la lutte contre le tabagisme. Celle-ci repose sur un ensemble de mesures législatives (loi Veil en 1976, loi Evin en 1991, décret de 2006 sur l'interdiction de fumer dans les lieux collectifs), de mesures fiscales et financières (taxations), de campagnes d'information (dispensées notamment par l'Institut National de Prévention et d'Education pour la Santé - INPES), d'actions locales et de services médicaux spécialisés avec consultations d'aide à l'arrêt (Almaric, 2007). Ce type de mesures, s'appuyant sur l'implication des professionnels de santé sur le terrain et sur une « auto-responsabilisation » de la population, semblent avoir en partie porté leur fruit, en particulier chez les hommes : la diminution de la consommation de tabac amorcée il y a quelques années serait à l'origine de la stabilisation voire de la diminution de l'incidence et de la mortalité observée récemment (Figure 14B). Les femmes restent, à l'inverse, particulièrement exposées aux méfaits de la cigarette ; elles sont devenues à la fois les cibles privilégiées des industriels du tabac et des campagnes récentes de prévention.

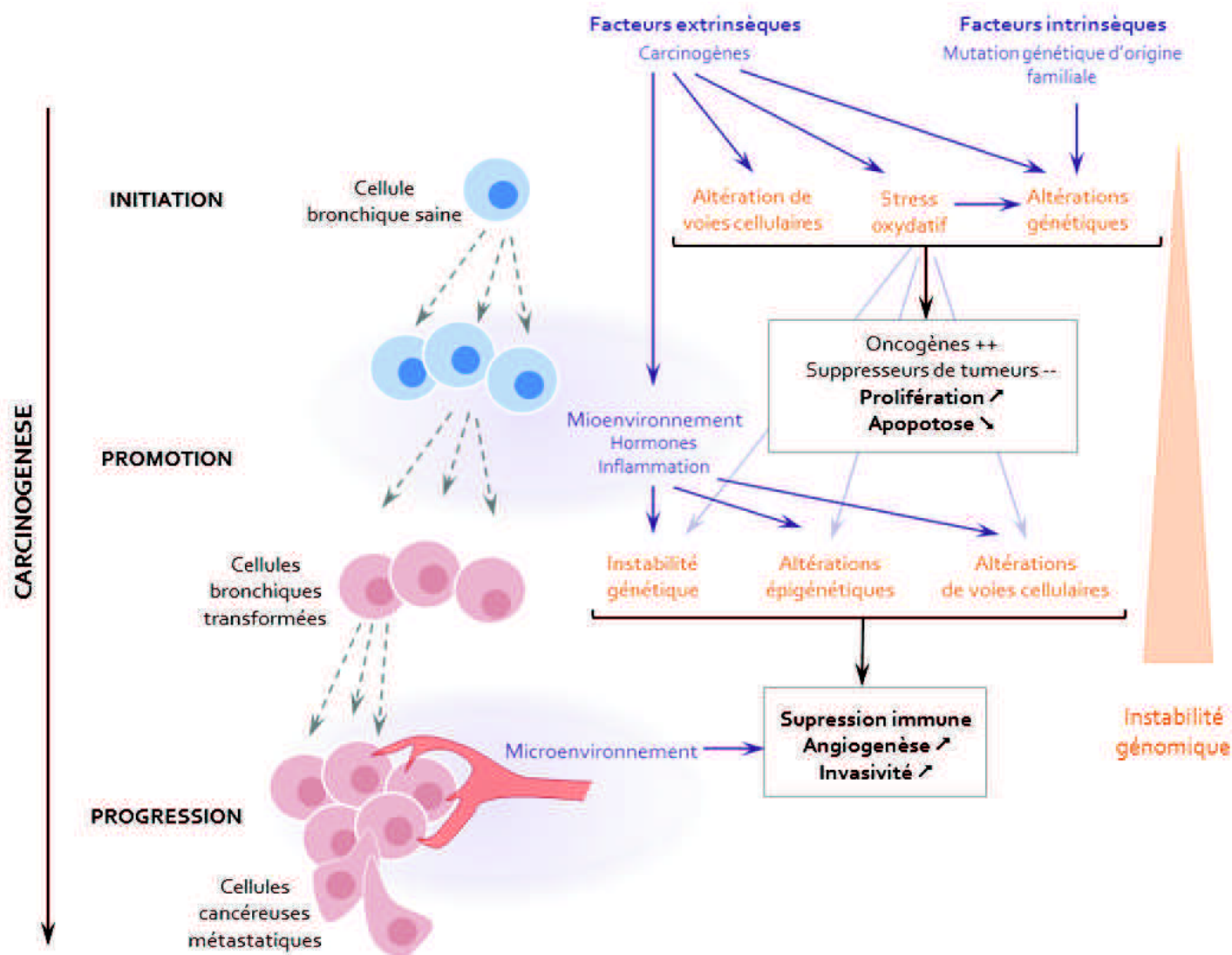


Figure 62 : La carcinogénèse pulmonaire, un processus multi-étapes, reflet des altérations qui s'accumulent dans le génome

Schématisation du caractère séquentiel de la carcinogénèse bronchique.

Au fur et à mesure de la transformation, les altérations génétiques et chromosomiques s'accumulent dans le génome des cellules de l'épithélium bronchique. Elles permettent l'émergence de nouvelles caractéristiques cellulaires qui sont sélectionnées par expansion clonale.

D'autre part, de nombreuses avancées ont été réalisées en recherche fondamentale pour décrypter les bases moléculaires du processus de carcinogenèse pulmonaire. Elles ont notamment permis de caractériser des dizaines d'oncogènes et de suppresseurs de tumeur impliqués dans le développement des carcinomes bronchiques. La vision actuelle de la carcinogenèse pulmonaire est celle d'un processus multi-étapes, s'étalant parfois sur des dizaines d'années, où les caractéristiques phénotypiques acquises par les cellules cancéreuses pulmonaires (prolifération accrue, résistance à l'apoptose, invasivité) sont le reflet des altérations - génétiques et épigénétiques - qui s'accumulent progressivement dans le génome, associées à l'influence du microenvironnement tumoral (**Figure 62**).

Les nouvelles méthodes de caractérisation moléculaire des tumeurs pulmonaires, et notamment les méthodes d'analyse génétique qui ont récemment culminé avec les premiers séquençages complets de génomes de tumeurs pulmonaires, ont permis d'asseoir le concept de profilage moléculaire des cancers du poumon. Ces approches ont mis en lumière la pertinence de la classification histologique des tumeurs pulmonaires, étant donné que les CBNPC et les CBPC présentent des profils d'altérations distincts (**Figure 63**) ([Kitamura et al., 2008](#)). Ainsi, on observe plus fréquemment dans les CBPC des altérations de *TP53*, *RB*, *c-MET*, *LKB1*, *PIK3CA* ou *BRAF*. Ils surexpriment souvent la protéine anti-apoptotique Bcl-2, mais aussi le récepteur c-Kit et son ligand, dans une boucle proliférative autocrine/paracrine quasiment jamais observée dans les CBNPC. A l'inverse, les CBNPC sont le siège d'une activation des récepteurs tyrosine-kinase HER-2 et EGF-R, de mutations de *KRAS* et d'inactivations de *CDKN2A*, rarement observées dans les CBPC ([Sato et al., 2007](#)). Cette caractérisation moléculaire permet également de préciser encore « l'identité » de chaque tumeur pulmonaire en décrivant précisément les aberrations qu'elle porte et les voies cellulaires qu'elle active. Cela a des implications importantes dans l'application de thérapies ciblées, adaptées à chaque tumeur, en prenant en compte ses particularités moléculaires ([Dacic, 2011](#)). Ainsi, des mutations dans le domaine tyrosine-kinase de l'EGF-R sont recherchées chez les patients CBNPC et constituent une indication de traitement ciblé par les inhibiteurs TK.

Malgré ces progrès, les traitements ne parviennent pas à limiter la mortalité due aux néoplasies pulmonaires : aucune amélioration de la survie n'a été obtenue au cours des dernières décennies et même les approches thérapeutiques les plus récentes et les plus ciblées démontrent leurs limites. On est ainsi confronté par exemple au problème de la résistance secondairement acquise, apparaissant après une réponse initiale, chez les patients mutés dans l'EGF-R et traités aux TKI ([Engelman et al., 2007](#) ; [Pao et al., 2005](#)). Dans ce contexte sombre, et compte tenu des limites actuelles de l'aide au sevrage tabagique, des délais d'efficacité de l'arrêt du tabac et des résultats globalement décevants des traitements des carcinomes bronchiques, la communauté scientifique est sans cesse à la recherche de nouvelles techniques préventives et thérapeutiques. Ces développements s'effectueront notamment dans les domaines des stratégies de détection précoce, d'identification de nouvelles cibles thérapeutiques, mais également d'identification de nouveaux facteurs de pronostic susceptibles d'orienter les stratégies de traitement.

La question du pronostic d'un patient atteint d'un cancer du poumon, et indirectement du temps de survie qui peut être espéré, est un point essentiel à la fois pour le patient et pour l'équipe soignante. En effet, des marqueurs de pronostic doivent permettre d'aider les cliniciens dans la prise en charge du malade et de sa famille, et les orienter vers le traitement approprié. Toutefois, évaluer le pronostic d'un patient atteint d'une tumeur pulmonaire est une tâche ardue et sensible. Une première indication est fournie par le type histologique de la tumeur. Ainsi, les tumeurs pulmonaires à petites cellules sont considérées comme les plus agressives. Présentant une croissance et une invasivité marquées, les CBPC sont souvent détectés à des stades avancés, avec fréquente présence de métastases. Ne pouvant le plus souvent pas donner lieu à une chirurgie du fait de la dissémination et considérés comme rarement chimiocurables, ces tumeurs ne laissent que peu de chance de survie aux patients qui en sont atteints, la médiane de survie des patients porteurs d'une forme disséminée ne dépassant pas 10 mois. Une deuxième indication, en particulier pour les CBNPC, vient de l'évaluation du stade de la tumeur grâce aux paramètres TNM. Plus le stade est élevé, plus le développement

Altérations moléculaires		CBPC	CBNPC
Oncogènes	Amplification de <i>ERB-B2</i> (HER-2)	Non	10 – 30%
	Mutation dans <i>ERB-B2</i> (HER-2)	ND	4%
	Amplification de <i>ErB-B1</i> (EGF-R)	Non	25 – 50%
	Mutation dans le domaine tyrosine-kinase de l'EGF-R	Non	10 – 40%
	Surexpression de c-KIT	40 – 70%	Non
	Mutation dans <i>RAS</i>	< 1%	15 – 25%
	Amplification de <i>MYC</i>	18 – 31%	8 – 20%
	Détection de Bcl-2 (IHC)	75 – 95%	10 – 35%
Gènes suppresseurs de tumeurs	LOH de <i>TP53</i> (p53)	90%	65%
	Mutation dans <i>TP53</i> (p53)	75%	50%
	LOH de <i>RB</i>	67%	31%
	Inactivation de RB	~100%	< 15%
	LOH de <i>CDKN2A</i> (p16-INK4A)	53%	66%
	Mutation dans <i>CDKN2A</i> (p16-INK4A)	< 1%	10 – 40%
	Hyperméthylation de <i>CDKN2A</i> (p16-INK4A)	~10%	~30%

Figure 63 : Patrons d'altérations moléculaires distincts entre CBPC et CBNPC

Les tumeurs pulmonaires, selon leur type histologique, n'exhibent pas avec la même fréquence les altérations moléculaires les plus courantes dans les carcinomes bronchiques. Ainsi, les altérations (amplifications, mutations activatrices) des récepteurs membranaires de la famille HER à rôle oncogène sont caractéristiques des tumeurs non à petites cellules. A l'inverse, les pertes de fonction des suppresseurs de tumeur p53 et Rb (perte d'homozygotie, mutations inactivatrices) sont plus fréquentes dans les cancers pulmonaires à petites cellules.

CBPC : carcinome bronchique à petites cellules ; CBNPC : carcinome bronchique non à petites cellules ; ND : non déterminé ; IHC : immunohistochimie ; LOH : perte d'homozygotie

Construit à partir de (Sato et al., 2007) et (Janku et al., 2010).

de la tumeur est avancé et plus la survie est menacée. Mais pour préciser davantage le pronostic d'un patient, les cliniciens ont depuis plus de vingt-cinq ans été à la recherche d'autres marqueurs de pronostic, en tentant de corréler des données cliniques sur le patient ou des paramètres intrinsèques de la tumeur avec la survie du patient, notamment par des analyses de survie menées sur de larges cohortes de patients (Osterlind and Andersen, 1986). Des paramètres sanguins, comme l'hypoalbuminémie, l'hyperuricémie, l'élévation des LDH et des phosphatases alcalines, l'anémie ou encore la thrombopénie ont depuis longtemps été considérées comme de marqueurs de mauvais pronostic (Tourani). Plus récemment, on a pu observer que de nombreuses altérations épigénétiques sont corrélées à des pronostics sombres. C'est notamment le cas de la surexpression des DNMT (Lin et al., 2007b), de la baisse d'expression des HDAC de classe II (Osada et al., 2004), de l'hyperméthylation de CDKN2A (Lou-Qian et al., 2013) ou de l'expression ectopique de plusieurs gènes CT (Gure et al., 2005).

Dans cette étude, j'ai identifié un gène dont l'expression au sein de la tumeur est associée à un pronostic sombre pour le patient. Ce gène a été identifié grâce à une approche innovante, basée sur une cohorte de 300 patients suivis pendant 10 ans, conciliant l'identification de la signature génique associée aux tumeurs agressives et la recherche d'expressions ectopiques de gènes tissu-spécifiques. Il s'agit du gène *PRL*, codant la prolactine.

3.2. L'expression ectopique du gène *PRL* dans les cellules pulmonaires cancéreuses

Notre travail sur l'expression du gène *PRL* dans le cancer du poumon s'est inscrit dans une mouvance assez récente d'intérêt croissant pour les sources extra-hypophysaires de prolactine, longtemps méconnues ou négligées (Ben-Jonathan et al., 1996). Depuis la découverte de l'hormone hypophysaire dans les années 1930 et pendant les soixante années qui ont suivi, les scientifiques se sont intéressés uniquement aux rôles de la prolactine hypophysaire en tant qu'hormone circulante, agissant sur des organes cibles après avoir emprunté la voie sanguine. Ce n'est qu'à partir des années 1990 que l'on a démontré la production de prolactine en plusieurs sites extra-hypophysaires (le sein, le placenta ou le testicule notamment). Ces productions extra-hypophysaires ont été sujet à débat. En effet, des observations de la protéine en immunohistochimie dans des tissus non-hypophysaires ne permettaient pas d'exclure qu'il s'agisse en réalité d'une perfusion de la prolactine circulante dans ces tissus, par infiltration de macrophages ou liaison à des protéines fixant la prolactine (PRL-binding proteins). Ce sont les observations de l'ARN messager qui ont mis fin à la controverse (Dagvadorj et al., 2007 ; Shaw-Bruha et al., 1997 ; Touraine et al., 1998). Le poumon est absent de toutes les études et revues répertoriant les sites de production extra-hypophysaires de prolactine (Bernichtein et al., 2010 ; Bole-Feysot et al., 1998 ; Ignacak et al., 2012). Les rares études qui se sont intéressées à la production de prolactine dans des panels tissus sains incluant le poumon s'accordent à dire qu'aucune production de prolactine n'a lieu dans le poumon. Ces données sont cohérentes avec les résultats que j'ai obtenus dans ce travail de thèse : je n'ai pu détecter aucune expression du gène *PRL* dans le poumon adulte sain.

Alors que le gène *PRL* voit son expression réprimée dans le poumon sain, j'ai détecté une expression de *PRL* à la fois dans des tumeurs pulmonaires et dans des lignées cellulaires de cancer du poumon. L'expression du gène *PRL* avait déjà été observée au cours des années 1990 dans les tumeurs du sein (Fields et al., 1993) et de la prostate (Nevalainen et al., 1997), mais c'est la première fois que la présence d'ARN messager *PRL* est démontrée dans les cellules cancéreuses pulmonaires. De plus, sein et prostate expriment déjà le gène *PRL* à un niveau basal dans un état non-pathologique. A l'inverse, dans les cellules pulmonaires cancéreuses, nous avons affaire à une réactivation d'un gène normalement silencieux, c'est-à-dire à une expression ectopique.

A.

		Patients	Patients <i>PRL</i> +	Proportion parmi les patients <i>PRL</i> +	Patients <i>PRL</i> -Like	Proportion parmi les patients <i>PRL</i> -Like
CBNPC	ADC	85	1	3,7%	0	
	CEB (+ BAS)	100	3	11,1%	3	8,8%
	CB(NE)GC	56	16	59,3%	16	47,1%
	CARCI	24	0		3	8,8%
CBPC		21	7	25,9%	12	35,3%
Autres		7	0		0	
TOTAL		293	27		34	

B.

		Patients	Patients <i>PRL</i> +	Proportion dans chaque type histologique	Patients <i>PRL</i> -Like	Proportion dans chaque type histologique	Patients <i>PRL</i> + et <i>PRL</i> -Like	Proportion dans chaque type histologique
CBNPC	ADC	85	1	1,2%	0		1	1,2%
	CEB (+ BAS)	100	3	3,0%	3	3,0%	6	6,0%
	CB(NE)GC	56	16	28,6%	16	28,6%	32	57,1%
	CARCI	24	0		3	12,5%	3	12,5%
CBPC		21	7	33,3%	12	57,1%	19	90,5%
Autres		7	0		0		0	
TOTAL		293	27	9,2%	34	11,6%	61	20,8%

C.

		Littérature	Cohorte
CBNPC	ADC	25-40%	29%
	CEB (+BAS)	20 - 30%	34%
	CB(NE)GC	10-15%	19%
	CARCI	<5%	8%
CBPC		15%	7%
Autres		5-10%	2%

Figure 64 : Fréquences de l'activation ectopique du gène *PRL* dans les tumeurs pulmonaires de la cohorte

A. Ce tableau illustre les proportions des différents types histologiques représentés parmi les patients *PRL* + et *PRL* - / *PRL* + like. Ainsi, plus de la moitié des patients *PRL* + souffrent de tumeurs neuroendocrines à grandes cellules.

B. Ce tableau illustre les proportions de patients *PRL* + et *PRL* - / *PRL* + like au sein des différents types histologiques. Ainsi, 28% des patients à tumeurs neuroendocrines à grandes cellules de la cohorte sont des patients *PRL* +. Plus de 90% des patients souffrant de carcinomes bronchiques à petites cellules sont considérés comme *PRL* + ou *PRL* - / *PRL* + like.

C. Ce tableau compare les proportions des différents types histologiques au sein de l'ensemble des tumeurs pulmonaires diagnostiquées (données de la littérature) et au sein de notre cohorte. Celle-ci est particulièrement enrichie en tumeurs épidermoïdes, en tumeurs neuroendocrines à grandes cellules et en tumeurs carcinoïdes.

CBPC : carcinomes bronchiques à petites cellules ; CBNPC : carcinomes bronchiques non à petites cellules ; ADC : adénocarcinomes ; CEB (+ BAS) : carcinomes épidermoïdes bronchiques et leur sous-type basaloïde ; CB(NE)GC : carcinomes bronchiques (neuroendocrines) à grandes cellules ; CARCI : tumeurs carcinoïdes.

Quelle est l'ampleur de cette expression hors contexte ? Dans notre étude, ce sont presque 10% des tumeurs pulmonaires de la cohorte que nous avons identifiées comme *PRL*-positives par une approche transcriptomique (**Figure 35B et 64B**). Cependant, cette approche nous a conduits à manquer des patients *PRL*+. En effet, certains patients *PRL*- en transcriptomique démontrent une amplification de la séquence de *PRL* en RT-PCR grâce à des primers situés dans les exons 4 et 5. Cette adéquation pas tout à fait parfaite entre données transcriptomiques et données de PCR a déjà pu être observée et discutée, notamment dans le cas des activations ectopiques de gènes CT dans les tumeurs pulmonaires où la technique de RT-PCR s'avère plus sensible que les données transcriptomiques (Rousseaux, 2013). Suite à l'observation d'un sous-groupe de patients *PRL*- / *PRL*+ like (considérés comme *PRL*-négatifs en transcriptomique mais présentant une signature génique très proche de celle des patients *PRL*+), nous avons suggéré que ce groupe de patients *PRL*- / *PRL*+ like puisse correspondre à des patients en réalité *PRL*+. Même si cette hypothèse, vérifiée uniquement sur un patient, n'a pas pu être validée à grande échelle, elle suggère que la proportion de patients réellement *PRL*+ puissent être plus élevée que les 10% estimés dans un premier temps (**Figure 64B**). D'autre part, il faut signaler que le chiffre de 10% de patients *PRL*+ dans notre cohorte doit être pris avec précaution et ne peut pas être extrapolé à l'ensemble des carcinomes bronchiques, étant donné que notre cohorte est biaisée vers certains types histologiques (**Figure 64C**). En effet, elle est particulièrement riche en tumeurs épidermoïdes, mais également en tumeurs neuroendocrines à grandes cellules qui expriment justement fréquemment le gène *PRL*.

Au sein de notre cohorte, nous avons noté un net enrichissement du groupe de tumeurs *PRL*-positives en tumeurs neuroendocrines (**Figure 64A**). En effet, près de 60% des tumeurs *PRL*+ sont des tumeurs neuroendocrines à grandes cellules et plus de 25% sont des carcinomes à petites cellules. Dans les premières classifications des tumeurs pulmonaires, les tumeurs neuroendocrines constituaient une entité à part. Aujourd'hui, elles apparaissent « diluées » au sein de la classification histopathologique. En effet, un phénotype de type neuroendocrine peut être observé principalement au sein des CBPC, mais également chez certains CBNPC (carcinomes neuroendocrines à grandes cellules et tumeurs carcinoïdes) (Linnoila, 1996). Au sein d'une tumeur, le caractère neuroendocrine se caractérise au niveau cellulaire par la présence de granules cytoplasmiques contenant des substances neuroendocrines. On pense que les tumeurs pulmonaires à caractéristiques neuroendocrines peuvent dériver d'un type cellulaire particulier, les cellules neuroendocrines, réparties au sein de l'épithélium bronchique. Celles-ci assurent dans le poumon sain un rôle qui se situe entre celui d'une cellule productrice d'hormone et une cellule nerveuse. Elles sont responsables de la sécrétion de peptides et hormones variés (sérotonine, bombésine, calcitonine) contrôlant de nombreuses fonctions au sein du tissu pulmonaire. Elles sont notamment responsables de la contraction ou du relâchement des muscles qui entourent les bronches, constituent le premier type histologique à apparaître dans le poumon fœtal et joueraient un rôle important dans la maturation des autres types cellulaires (Siegfried et al., 1994). De nombreux peptides et hormones neuroendocrines ont été décrits comme sécrétés par les tumeurs pulmonaires neuroendocrines (et en particulier par les CBPC). On peut y retrouver l'hormone anti-diurétique (ADH), l'hormone corticotrope (ACTH), l'arginine vasopressine (AVP) ou le facteur natriurétique auriculaire (ANP) (Campling et al., 1995 ; Siegfried et al., 1994). Cette sécrétion inappropriée d'hormones par les cellules cancéreuses conduit souvent à des syndromes paranéoplasiques endocriniens (Rosti et al., 2006), tels que le syndrome paranéoplasique de Schwartz-Bartter (dû à la sécrétion d'ADH) ou le syndrome cushingoïde (sécrétion d'ACTH). Il est intéressant de relever que l'hormone de croissance (growth hormone GH), l'hormone folliculo-stimulante (FSH), l'hormone lutéinisante (LH) ou la thyroestimuline (TSH) ne sont jamais ou exceptionnellement sécrétées par des cellules cancéreuses pulmonaires, peut-être du fait de la nécessité d'exprimer deux sous-unités différentes, de les glycosyler et de les assembler pour obtenir une hormone active (Gropp et al., 1980). La sécrétion de prolactine par des cellules cancéreuses pulmonaires n'a jamais été formellement montrée. Dans l'ensemble de la littérature, seul un papier datant des années 1980, décrivant une grande variété d'hormones détectées dans les milieux de culture d'une quinzaine de lignées de CBPC, a décrit une faible concentration de prolactine dans le surnageant de culture d'une lignée (Sorenson et al., 1981).

Dans notre étude, tumeurs à petites cellules et tumeurs neuroendocrines à grandes cellules constituent la majeure partie des tumeurs *PRL*-positives (**Figure 64A**). En cohérence avec cette observation sur patients, c'est au sein de 3 lignées cellulaires de CBPC (H146, H524 et H69) que j'ai identifié une expression ectopique du gène *PRL*, alors qu'aucune expression n'a pu être détectée dans les lignées de CBNPC testées (H1299 ou H1650 par exemple). Nous pouvons toutefois regretter le fait de n'avoir pas disposé de lignées cellulaires de cancer neuroendocrine à grandes cellules *PRL*-positives qui auraient pu constituer de bons modèles complémentaires dans l'étude de l'expression de *PRL*. Notre étude a démontré de manière inédite l'expression ectopique du gène *PRL* dans les cellules cancéreuses pulmonaires. Dès lors, la question s'est posée de l'éventuelle contribution de l'expression de ce gène codant la prolactine dans la tumorigenèse pulmonaire. Et nous allons le voir, l'histoire de la prolactine et du cancer est une histoire ancienne et complexe.

3.3. Prolactine et cancers : une histoire à deux visages

La prolactine est une hormone hypophysaire qui promeut le développement de la glande mammaire, durant la puberté et au cours d'une grossesse (**Figure 32**). Etant donné le rôle de la prolactine dans la prolifération et la survie des cellules de l'épithélium mammaire, la question de l'implication de la prolactine dans la tumorigenèse humaine, cancer du sein en particulier, s'est rapidement posée, a longtemps été débattue et reste un champ de recherche très actif ([Fernandez et al., 2010](#)). Au début des années 2000 ont émergés deux mécanismes distincts d'action de la prolactine sur la croissance tumorale. Le premier repose sur l'action de la prolactine circulante, alors que le second implique la prolactine produite localement ([Ben-Jonathan et al., 2002](#)) (**Figure 65A et B**).

3.3.1. Prolactine circulante et cancer du sein

Le rôle de la prolactine dans les néoplasies mammaires a été intensément débattu. On a observé dès les années 1970 que la prolactine en tant qu'hormone contribuait à la pathogenèse et la progression des tumeurs mammaires chez la souris ([Welsch and Nagasawa, 1977](#)). Mais, au cours des années 1980, les études épidémiologiques n'ont pas permis d'asseoir cette relation chez l'humain ([Tworoger and Hankinson, 2008](#)). Certaines études ont mis en évidence des niveaux sériques de prolactine plus élevés chez les femmes avec une histoire familiale de cancers du sein ([Levin and Malarkey, 1981](#) ; [Love and Rose, 1985](#)). Mais ces résultats ont été contredits par d'autres travaux ([Fishman et al., 1978](#) ; [Ingram et al., 1990](#)), ces conclusions différentes pouvant être dues aux faibles tailles des cohortes, à la fiabilité limitée des dosages sériques ou aux variations dans les phases du cycle menstruel ou le statut (ménopausées ou non) des femmes testées ([Koenig et al., 1993](#)). D'autre part, les essais cliniques menés chez des patientes souffrant de cancer du sein traitées avec des agents pharmaceutiques inhibant la synthèse hypophysaire de prolactine (bromocriptine notamment) se sont révélés être des échecs ([Anderson et al., 1993](#) ; [Bonnetterre et al., 1988](#)). Ce n'est que bien plus tard, à la fin des années 1990, que de nouvelles études prospectives ont ré-établi le rôle de la prolactine en tant qu'hormone dans la tumorigenèse mammaire. L'une de ces études, publiée en 1999, a été réalisée sur une très large cohorte d'infirmières (plus de 32 000 individus). Des échantillons sanguins ont été prélevés entre 1989 et 1990 puis la cohorte été suivie durant 5 années pendant lesquelles 306 cas de cancer du sein se sont déclarés ([Hankinson et al., 1999](#)). Ce travail a permis d'observer une association positive statistiquement significative entre le niveau sérique de *PRL* et le risque de développer un cancer du sein (**Figure 65A**) : les femmes dans le dernier quartile (25% des prolactinémies les plus élevées) ont deux fois plus de risque de développer une tumeur mammaire que les femmes du premier quartile. Des études ont également démontré un lien entre les niveaux sériques de prolactine et le pronostic des cancers du sein, une concentration sanguine élevée étant associée à une survie diminuée ([Wang et al., 1995](#)) et à une plus grande résistance aux traitements ([Bhatavdekar et al., 1994](#)).

3.3.2. Prolactine locale et cancer du sein

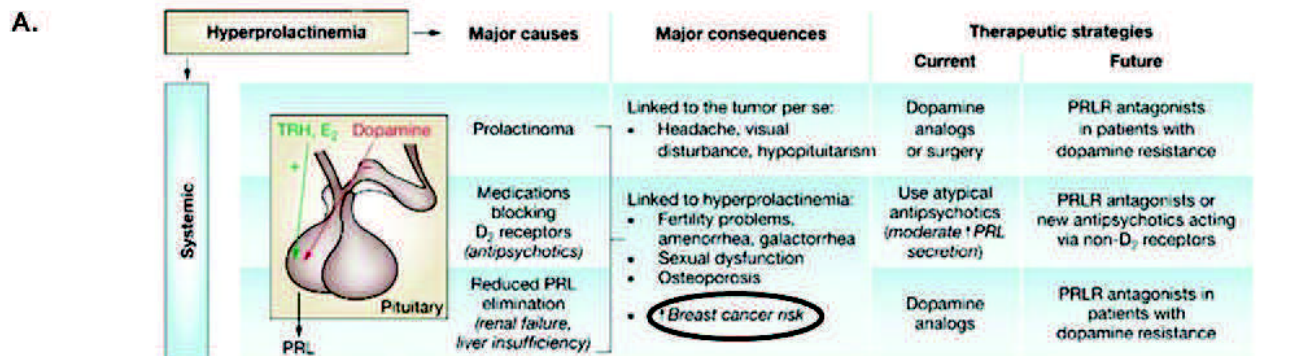
Durant le temps qu'a duré la polémique sur l'implication de l'hormone dans les cancers du sein, la communauté scientifique a commencé remettre en question le dogme de la prolactine agissant uniquement en tant qu'hormone circulante (Bernichtein et al., 2010). Au début des années 1990, on détecte en effet l'expression du gène *PRL* dans les tissus mammaires sains et cancéreux (Fields et al., 1993). Cela pousse les scientifiques à considérer la prolactine également en tant que facteur de croissance, produit localement et agissant localement dans le tissu mammaire pour promouvoir la prolifération des cellules de l'épithélium. Cette sécrétion locale expliquerait notamment les échecs des traitements par la bromocriptine, agoniste dopaminergique, qui n'est pas capable d'inhiber la sécrétion par les sites de production extra-hypophysaires.

Les cellules cancéreuses mammaires en culture produisent des quantités non négligeables de prolactine dans le milieu (Reynolds et al., 1997). D'autre part, plus de 98% des tumeurs mammaires expriment le récepteur de la prolactine sous différentes isoformes. On a même observé une surexpression du récepteur dans la tumeur comparé aux tissus mammaires sains adjacents (Figure 65A) (Touraine et al., 1998), surexpression qui pourrait être due à une amplification du gène *PRL-R* (Tran-Thanh et al., 2011). On fait rapidement l'hypothèse d'une boucle autocrine/paracrine au sein de la tumeur qui favoriserait la croissance tumorale (Clevenger et al., 1995). Cette hypothèse a été confortée en modèle murin, puisque la surexpression dirigée de la prolactine dans la glande mammaire conduit au développement de carcinomes invasifs (Arendt et al., 2011 ; Rose-Hellekant et al., 2003) et qu'à l'inverse, la perte de l'expression du récepteur ralentit la croissance tumorale (Oakes et al., 2007).

On a depuis confirmé chez l'Homme que l'activation de PRL-R à la surface des cellules cancéreuses stimule la prolifération, inhibe la différenciation, favorise l'acquisition de l'invasivité et augmente la résistance aux traitements des cellules cancéreuses mammaires (Figure 65A) (Clevenger et al., 2003). L'effet de la prolactine sur la prolifération des cellules cancéreuses mammaires passe par des dérégulations de protéines du cycle (Schroeder et al., 2002), dont la cycline D1 qui est surexprimée dans plus de 50% des cancers du sein. L'effet mitogène serait médié non seulement par la voie JAK/STAT mais également par une activation de la voie ERB/Ras/Raf/MAPK (Yamauchi et al., 2000). L'effet sur la survie impliquerait notamment la voie PI3K/Akt (Neilson et al., 2007).

3.3.3. Prolactine locale et autres cancers

Etant donné que la prolactine locale semble favoriser la transformation des cellules mammaires, et que la prolactine est également produite d'autres tissus (prostate, utérus, ovaire), on a cherché plus récemment à évaluer l'implication de la prolactine sécrétée localement dans le développement de cancers variés (Sethi et al., 2012). Le rôle de la prolactine a été investigué dans les cancers de l'ovaire et de l'endomètre qui surexpriment fréquemment le récepteur (Levina et al., 2009). La boucle autocrine/paracrine semble participer à la progression tumorale, en conjonction avec la prolactine circulante (Clendenen et al., 2013). Mais c'est dans le cancer de la prostate que le plus haut niveau de compréhension a été atteint, et qu'on a démontré l'implication de la prolactine, ce « joueur féminin », dans un « jeu d'hommes » au sein de cet organe mâle par excellence (Goffin et al., 2011). Les données épidémiologiques ne démontraient pas de corrélation entre cancer de la prostate et niveau de prolactine circulante (Stattin et al., 2001) alors que les preuves se sont accumulées pour un rôle de la prolactine sécrétée localement (Figure 65B). Les cellules malignes de prostate produisent de la prolactine à un niveau élevé (Nevalainen et al., 1997). La prolactine induit la prolifération et inhibe l'apoptose de cellules prostatiques humaines en culture. L'expression dirigée de prolactine dans la prostate de souris conduit à l'apparition de lésions néoplasiques, voire d'adénocarcinomes (Rouet et al., 2010). La plupart des effets de la prolactine sur les cellules de la prostate semblent être comparables aux effets sur les cellules mammaires, et passer par la liaison de la prolactine sécrétée sur son récepteur membranaire et l'activation de la voie JAK/STAT (Dagvadorj et al., 2007).



Breast cancer relative risk

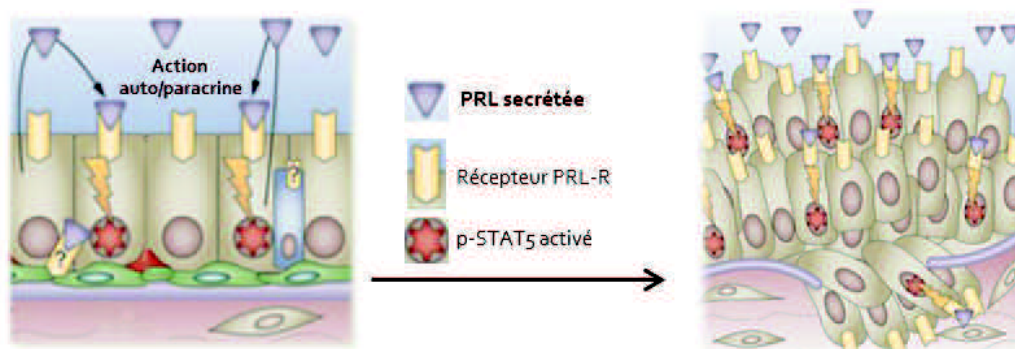
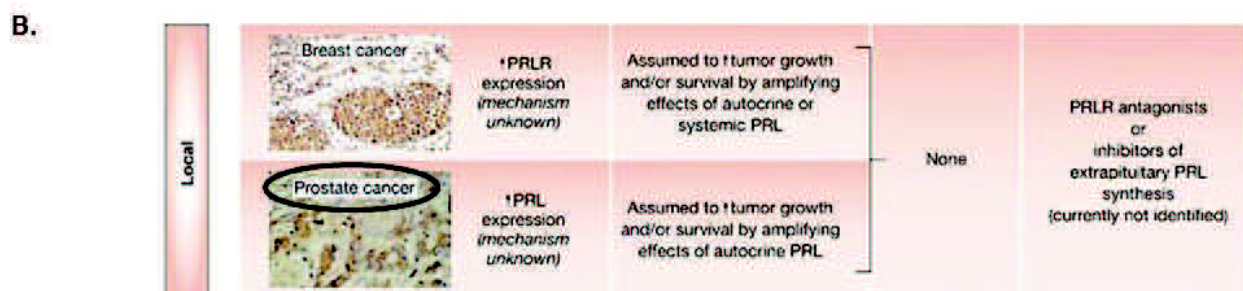
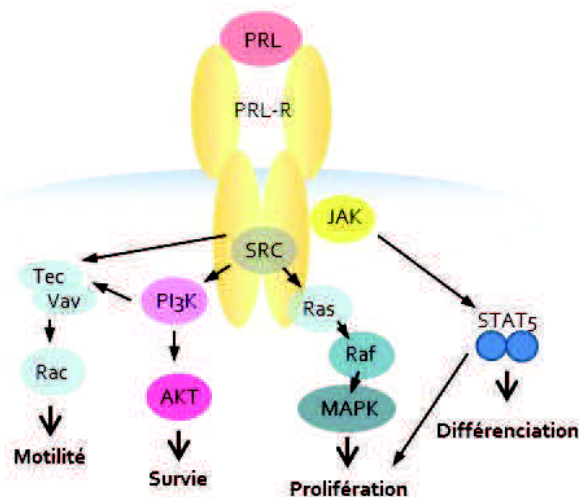
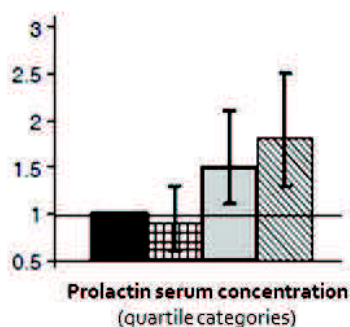


Figure 65 : Prolactine et cancer

A. Hyperprolactinémie systémique. Tiré de (Goffin et al., 2006).

Les femmes présentant une concentration sérique de prolactine dans le quartile le plus haut ont presque deux fois plus de risques de développer une tumeur mammaire. Tiré de (Tworoger et al., 2005).

Les voies cellulaires en aval du récepteur PRL-R sont impliquées dans l'acquisition du phénotype cancéreux des cellules mammaires. Adapté de (Clevenger et al., 2003) et de (Damiano and Wasserman, 2013).

B. La production locale de prolactine est impliquée dans la progression des tumeurs du sein et de la prostate. Tiré de (Goffin et al., 2006).

La sécrétion de la prolactine par les cellules de l'épithélium prostatique entraîne une boucle autocrine/paracrine qui promeut la prolifération des cellules. Tiré de (Goffin et al., 2011).

Puisque la production locale de prolactine ne peut être contrôlée par les agonistes dopaminergiques qui agissent sur la sécrétion hypophysaire uniquement, on cherche à développer des thérapies innovantes dans le traitement des cancers du sein et de la prostate notamment, basées sur des inhibiteurs du récepteur à la prolactine. L'antagoniste G129R-hPRL, qui bloque sélectivement la dimérisation de PRL-R, a d'ores et déjà démontré une capacité à induire l'apoptose de cellules cancéreuses mammaires (Chen et al., 1999). Quant à l'anticorps neutralisant LFA102 qui abroge la fonction du PRL-R *in vivo*, il démontre des effets anti-tumeurs prometteurs dans les modèles précliniques de cancers du sein et de la prostate (Damiano and Wasserman, 2013).

En cohérence avec ces observations du rôle de la prolactine locale dans les cancers du sein et de la prostate, nous avons adopté dans un premier temps l'hypothèse d'une sécrétion de la prolactine dans le milieu extracellulaire par les cellules tumorales pulmonaires et d'une action de cette protéine sécrétée localement dans une boucle autocrine/paracrine au sein de la tumeur. Mais nos observations ont rapidement conduit à remettre cette hypothèse en question.

3.4. Les cellules pulmonaires cancéreuses produisent un variant d'ARN messager *PRL* amputé des premiers exons

Le gène *PRL* possède deux promoteurs alternatifs à spécificité tissulaire. Dans l'hypophyse, la liaison du facteur de transcription Pit-1 sur le promoteur proximal du gène *PRL* aboutit à la production du variant 1 de l'ARN messager contenant l'exon 1 (Figure 68). Dans le placenta, le sein ou les lymphocytes, c'est le promoteur distal du gène *PRL* qui est utilisé, conduisant à la synthèse du variant 2 de l'ARN messager qui intègre la séquence de l'exon alternatif 1a situé en amont ainsi qu'une partie de l'exon 1 (DiMattia et al., 1988). C'est ce même promoteur distal qui dirige l'expression locale de la prolactine dans les tumeurs mammaires (Shaw-Bruha et al., 1997) (Figure 68), alors que les cellules malignes de prostate semblent utiliser les deux promoteurs, indépendamment de Pit-1, pour promouvoir une forte production de prolactine au sein de la tumeur (Dagvadorj et al., 2007). Les deux variants d'ARNm *PRL* ne diffèrent que dans leur 5'UTR, ce qui pourrait participer à une régulation tissulaire de la stabilité ou de la traduction de l'ARNm, mais leur traduction aboutit à la synthèse de la même protéine prolactine.

Dans le modèle du cancer du poumon, notre première surprise est venue du fait qu'aucun de ces deux promoteurs n'est utilisé dans les cellules cancéreuses pulmonaires. En effet, on est incapable d'amplifier en PCR les ADNc des lignées de CBPC *PRL*-positives avec des primers reconnaissant les 5'UTR des variants 1 et 2 de l'ARNm *PRL*. En d'autre terme, aucune des séquences 5'UTR codées par les exons 1a et 1 ne sont présentes dans les ARNm *PRL* produits par nos lignées de cancer du poumon. Puis des cartographies des ARNm *PRL* en PCR, sur lignées cellulaires mais aussi sur tumeurs de patients, nous ont démontré que les séquences absentes ne se limitaient pas au 5'UTR. Dans les lignées CBPC, exon 1a et 1 en entier (5'UTR et début de la séquence codante) ne sont pas transcrits, seul les séquences des exons 2 à 5 peuvent être détectées dans l'ARNm *PRL*. Chez les tumeurs *PRL*-positives, on observe toujours une absence des séquences des exons 1a et 1, mais les séquences des exons 2 et 3 peuvent parfois ne pas être détectées non plus.

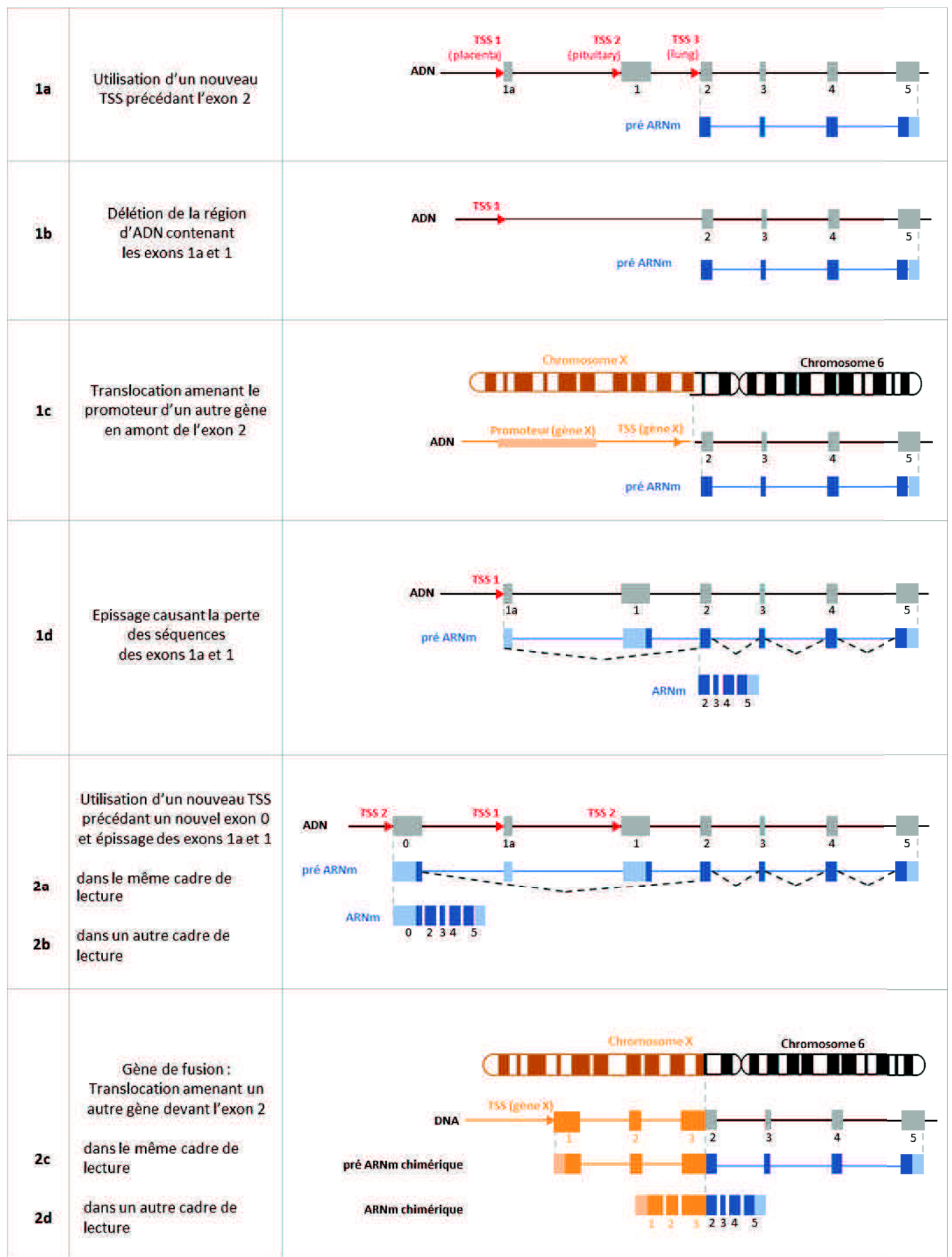


Figure 66 : Comment expliquer l'absence des premiers exons dans l'ARNm *PRL* produit par les cellules cancéreuses pulmonaires ?

Nous pouvons alors émettre des hypothèses variées pour tenter d'expliquer l'absence des séquences des premiers exons dans les transcrits produits par les cellules cancéreuses pulmonaires (**Figure 66**). On peut tout d'abord envisager l'utilisation d'un nouveau signal de début de transcription (TSS), encore jamais décrit, suivant l'exon 1, et ne conduisant au début de transcription qu'à partir de l'exon 2 (**Hypothèse 1a**). On peut aussi imaginer que le locus 6p22 ait subi une délétion génomique hétérozygote conduisant à la disparition des exons 1a et 1 du gène *PRL* (**Hypothèse 1b**). Un autre évènement chromosomique, de type translocation, pourrait amener les séquences régulatrices d'un autre gène en amont de l'exon 2 du gène *PRL* et conduire à une transcription uniquement à partir de cet exon (**Hypothèse 1c**). On peut encore penser à l'existence d'un épissage aberrant qui supprimerait les premiers exons de l'ARN messenger *PRL* produit par les cellules cancéreuses pulmonaires (**Hypothèse 1d**). En plus de son promoteur distal extra-hypophysaire, le gène *PRL* pourrait posséder un nouveau promoteur encore plus en amont, précédant un nouvel exon O. La synthèse du transcrit long serait suivi par l'épissage des exons 1a et 1. Dans ce cas, le début de traduction dans l'exon O pourrait se faire dans le même cadre de lecture que la prolactine (**Hypothèse 2a**) ou dans un autre cadre (**Hypothèse 2b**). Enfin, nous avons imaginé le cas où une translocation chromosomique amènerait en amont du gène *PRL* un autre gène et son promoteur, conduisant à la synthèse d'un ARN messenger chimérique et d'une protéine de fusion, par traduction soit dans le cadre de lecture de la prolactine (**Hypothèse 2c**) soit dans un autre cadre (**Hypothèse 2d**).

Dans les quatre cas notés (**1**), on obtiendrait un ARN messenger court, d'une longueur maximale de 824 pb au lieu des 1043 pb du variant placentaire (si uniquement l'exon 1 est manquant). Cet ARN messenger tronqué serait dépourvu de son codon ATG de début de traduction. Nous nous sommes demandés si la traduction pourrait quand même avoir lieu en utilisant un autre codon start en aval (**Figure 67A**). Il existe 13 triplets ATG dans le reste de la séquence de l'ARNm *PRL*. Ce codon doit être dans un environnement particulier, appelé séquence Kozak (de consensus gccRccATGG) pour être possiblement utilisé comme début de traduction par le ribosome. J'ai alors utilisé l'outil « NetStart 1.0 Prediction Server » qui affecte un score à chaque ATG pour déterminer s'il peut potentiellement correspondre à un signal de début de traduction. Selon cette analyse, 5 codons ATG en aval pourraient être utilisés pour initier la transcription. Ces 5 ATG alternatifs sont tous situés dans le même cadre de lecture que la prolactine, et conduiraient après traduction à la production d'une prolactine tronquée : la protéine perdrait au minimum ses 63 premiers acides aminés.

A l'inverse, dans les cas notés (**2**), l'ARN messenger produit serait plus long (>1043 pb), car prolongé en 5' par une séquence inconnue (d'un exon O ou d'une séquence d'un autre gène fusionné). Si l'ajout de la séquence en 5' conserve le cadre de lecture (**2a et 2c**), la prolactine produite par traduction serait allongée à son extrémité amino-terminale par un nouveau « début » de séquence de taille inconnue. De plus, la prolactine ne serait pas entière : il lui manquerait au moins les 10 premiers acides aminés de son peptide signal codés par l'exon 1 manquant. Si l'ajout d'une séquence en 5' ne se fait pas dans la phase de lecture de la prolactine (**2b et 2d**), alors une protéine totalement différente pourrait être produite. La question se pose alors de savoir si dans la séquence de *PRL*, un autre cadre de lecture pourrait être assez long pour être utilisé (**Figure 67B**). On constate que les deux autres cadres de lecture de la séquence *PRL* présentent des codons stop précoces, après 31 et 37 aa seulement. L'utilisation de l'un de ses cadres de lecture ne permettrait probablement pas la synthèse d'une protéine, étant donné que la cellule possède des mécanismes de détection et de dégradation des ARN messagers non-sens présentant des codons stop trop précoces (nonsense mediated decay ou NMD) ([Chang et al., 2007](#)). L'observation par Northern blot d'un ARN messenger *PRL* chez les cellules CBPC d'une taille proche du transcrit placentaire (environ 800 pb) tend à exclure la possibilité de production d'un ARN messenger de grande taille, contenant des exons surnuméraires en 5' ou en fusion avec un autre gène (**Hypothèses 2**). Nous pensons donc qu'un ARNm *PRL* tronqué, amputé de ces premiers exons est produit selon l'un des mécanismes précédemment proposés (**Hypothèses 1**).

A.

Position	-6	-5	-4	-3	-2	-1	1	2	3	4	5	6	7	Score	Valide	Séquence de la protéine	Longueur de la protéine
																MNIKGSFWKGSLLLLLVSNLLCCSVAPLPICPGGAARCOVTLRDL FDRAVVLSHYIHNLSSSEMFSEFDKRYTHGRGFITKAINSCHTSSLAT PEDKEQAQQMNQKDFLSLIVSILRSWNEPLYHLVTEVRGMQEAPE EAILSKAVEIEEQTKRLLEGMEIVSQVHPETKENIYYPVWSGLPSL QMADEESRLSAYYNLLHCLRRDSHKIDNYLKLLKCRIHNNNC	227
114	T	G	C	C	C	G	A	T	G	C	C	A	G	0,411	no		
190	T	C	A	G	A	A	A	T	G	T	T	C	A	0,635	yes	MFSEFDKRYTHGRGFITKAINSCHTSSLATPEDKEQAQQMNQKDF LSLIVSILRSWNEPLYHLVTEVRGMQEAPEAILSKAVEIEEQTKRL EGMEIVSQVHPETKENIYYPVWSGLPSLQMADEESRLSAYYNLL HCLRRDSHKIDNYLKLLKCRIHNNNC	164
221	A	T	A	C	C	C	A	T	G	G	C	C	G	0,489	no		
307	C	A	A	C	A	G	A	T	G	A	A	T	C	0,679	yes	MNQKDFLSLIVSILRSWNEPLYHLVTEVRGMQEAPEAILSKAVEIEE QTKRLLEGMEIVSQVHPETKENIYYPVWSGLPSLQMADEESRLS AYYNLLHCLRRDSHKIDNYLKLLKCRIHNNNC	125
359	C	C	T	G	G	A	A	T	G	A	G	C	C	0,404	no		
397	C	G	T	G	G	T	A	T	G	C	A	A	G	0,593	yes	MOEAPEAILSKAVEIEEQTKRLLEGMEIVSQVHPETKENIYYPVW SGLPSLQMADEESRLSAYYNLLHCLRRDSHKIDNYLKLLKCRIHN NNC	95
472	G	A	G	G	G	C	A	T	G	G	A	G	C	0,546	yes	MELIVSQVHPETKENIYYPVWSGLPSLQMADEESRLSAYYNLLHC LRRDSHKIDNYLKLLKCRIHNNNC	70
515	A	A	G	A	A	A	A	T	G	A	G	A	T	0,391	no		
556	C	T	G	C	A	G	A	T	G	G	C	T	G	0,637	yes	MADEESRLSAYYNLLHCLRRDSHKIDNYLKLLKCRIHNNNC	42
563	T	G	G	C	T	G	A	T	G	A	A	G	A	0,351	no		
723	T	C	C	T	T	A	A	T	G	A	T	C	C	0,252	no		
772	G	A	A	T	C	C	A	T	G	C	T	T	G	0,068	no		
840	T	C	T	A	A	A	A	T	G	A	A	A	A	0,064	no		

B. Cadre de lecture 1 (Cadre de lecture utilisé pour la traduction de la protéine prolactine) : 217 acides aminés

SLLLLLVSNLLCCSVAPLPICPGGAARCOVTLRDLFDRAVVLSHYIHNLSSSEMetFSEFDKRYTHGRGFITKAINSCH
TSSLATPEDKEQAQQMMetNQKDFLSLIVSILRSWNEPLYHLVTEVRGMQEAPEAILSKAVEIEEQTKRLLEGMEIVSQVHPET
KENIYYPVWSGLPSLQMADEESRLSAYYNLLHCLRRDSHKIDNYLKLLKCRIHNNNCStopAHIHFIYF
StopEGPStopStopSVPLQASFSFSCISFESMetLGCNRSPLKKStopKLTStopRHQNLKStopKKK

Cadre de lecture 2 : 31 acides aminés

GPSCCWCQTCSCARAWPPCPSVPAGLPDARStopPFETCLTAPSSCPTTSITSPQKCSANSINGIPMetAGGSLPRPS
TAATLLPLPPKTRSKPNRStopIKKTFStopAStopStopSAYCDPGMetSLCIIWSRKYVVCKKPRRLSYPKLStopRLRSKPN
GFStopRAWSStopStopSARFILKPKKMetRSTLSGRDFHPCRWLMetKSLAFLLIITCSTAYAGIHKSTIIS55StopSAESST
TTAKPTSISSISEKVLNDPFHCKLLLVSLNPNCLGVTGLLLKNKNStopLLRDIKStopNEKK

Cadre de lecture 3 : 37 acides aminés

VPPAAAGVKPAPVPERGPLAHLSSRRGCPMetPGDPSRPVStopPRRRPVPLHPStopPLLRNVQRIRStopTVYPWPVGVHY
QGHQQLPHFFPCHPRROGASPTDESKRLSEPD5QHIAILEStopASVSSGHGSTWYARSPGGYPIQ5CRDStopGANQT
ASRGHGADSQPG55StopNQRKStopDLPCLVGT5IPADGStopStopRVSPFCLLStopPAPLPTQGF5StopNRQL5QAPEV
PNHPQQQLLSPHPFHLFLRRSLMetIR5IASFFStopLYLFStopIHAWVStopQV55StopKIKTD5LET5K5KMetKKK

Figure 67 : Autres signaux de début de traduction et cadres de lecture alternatifs dans l'ARNm *PRL* amputé des premiers exons

A. L'outil « NetStart 1.0 Prediction Server », développé par la Technical University of Denmark (DTU), est disponible en ligne (<http://www.cbs.dtu.dk/services/NetStart/>). Il calcule un score pour chaque codon ATG présent dans une séquence, selon la conservation de la séquence Kozak environnante. Quand le score est considéré comme valide, la séquence de la protéine produite lors de la traduction est indiquée.

B. Séquences des protéines qui seraient produites par traduction de l'ARNm *PRL* amputé de son exon 1, dans les trois cadres de lecture.

Pour tester ces différentes hypothèses, nous pouvons proposer différentes stratégies expérimentales. La présence d'une délétion génomique en amont du gène *PRL* pourrait être démontrée par une stratégie de PCR en amont de l'exon 2 sur ADN génomique de lignées CBPC. L'existence d'une translocation peut être recherchée par la technique d'hybridation *in situ* en fluorescence (Fluorescence In Situ Hybridization – FISH) grâce à des sondes établies sur BACs dont les séquences sont situées au niveau du locus *PRL* et en amont. Des éléments de réponse pourraient également être apportés par un RNAseq sur les lignées CBPC.

Plus généralement, il faut imaginer un évènement qui soit suffisamment variable pour expliquer les différentes étendues d'absences observées (juste l'exon 1 seulement chez le patient n°59 et dans les lignées CBPC, potentiellement jusqu'à l'exon 3 chez le patient n°61). Ainsi, l'hypothèse d'un TSS inconnu en amont de l'exon 2 doit être revue car elle ne pourrait pas expliquer les absences des séquences de l'exon 2 voire 3 observées chez certains patients. Malgré tout, cet évènement doit également être suffisamment courant. En effet, dans tous les cas de lignées cellulaires et de patients *PRL*-positifs que nous avons testés, nous avons observé ces absences des premiers exons. A moins que plusieurs mécanismes différents se combinent pour expliquer la production de différents variants ARNm *PRL* par les cellules pulmonaires cancéreuses *PRL*-positives.

Dans tous les cas, la synthèse d'un ARNm *PRL* déplété des séquences des premiers exons pourrait avoir des conséquences sur l'éventuelle prolactine produite par traduction. La perte de l'exon 1 peut conduire à la production d'une prolactine amputée de ses acides aminés amino-terminaux, acides aminés qui forment un peptide signal. Ainsi, la perte du peptide signal pourrait conduire à une prolactine incapable d'être adressée au réticulum et à l'appareil de Golgi, non-sécrétée dans le milieu extérieur et donc retenue dans la cellule. Ceci est en accord avec l'hypothèse d'une rétention *in cellulo* de la prolactine dans les cellules pulmonaires cancéreuses que nous avons émise après avoir observé l'absence de récepteur à la prolactine dans ces cellules. D'autre part, la perte d'une partie de la protéine pourrait expliquer pourquoi nous n'avons pas été en mesure de détecter la prolactine dans les cellules cancéreuses pulmonaires avec un anticorps potentiellement incapable de reconnaître cette forme tronquée de la protéine. Mais puisque je n'ai pour l'heure pas été capable de démontrer la présence de la protéine, on ne peut pas exclure la possibilité que cet ARNm *PRL* produit pas les cellules cancéreuses pulmonaires ne soit pas traduit. Malgré tout, nous avons démontré que la seule activation ectopique du gène *PRL* conduisant à la synthèse d'un ARN messenger est responsable d'une résistance accrue des cellules *PRL*-positives.

3.5. La résistance *PRL*-associée

Chez les cellules CBPC *PRL*+, une simple extinction de l'expression du gène *PRL* par un ciblage de l'ARNm *PRL* avec un siARN cause une augmentation significative de la mort cellulaire induite par un stress génotoxique. Ceci démontre l'implication directe de l'expression du gène *PRL*, sinon de la protéine *PRL*, dans l'acquisition par les cellules pulmonaires cancéreuses *PRL*+ d'une résistance accrue. Ce résultat nous confère des bases au phénotype agressif des tumeurs *PRL*+. Nous avons en effet démontré que les tumeurs qui expriment de manière aberrante le gène *PRL* sont des tumeurs agressives, qui sont associées à un pronostic sombre, car causant une diminution significative de la survie des patients correspondants. Nous proposons alors que cette agressivité de la tumeur *PRL*+ puisse être directement causée par l'acquisition par les cellules tumorales d'une plus grande résistance aux différents stress qu'elles subissent, parmi lesquels l'hypoxie, le manque de nutriments, la surveillance immune ou encore les traitements par des agents chimiothérapeutiques. De manière intéressante, les analyses transcriptomiques menées sur ces mêmes cellules CBPC (traitées ou non par siARN *PRL*) mais également sur les tumeurs de la cohorte font ressortir des gènes impliqués dans la résistance à l'hypoxie et aux privations ainsi que des gènes impliqués dans la résistance aux traitements anti-cancéreux.

Nous avons également identifié dans les analyses transcriptomiques des gènes impliqués dans le cycle cellulaire et la mitose, ainsi que dans la migration et l'adhérence. Ces observations sont dignes d'intérêt, car l'acquisition d'une prolifération et/ou d'une capacité invasive accrues par les cellules cancéreuses qui expriment

le gène *PRL* pourrait également expliquer l'agressivité des tumeurs *PRL*+. Néanmoins, dans nos modèles de cellules CBPC, nous n'avons observé aucun phénotype lié à la prolifération cellulaire après silençage du gène *PRL*. Il faut toutefois faire remarquer que ces analyses ont été réalisées dans une approche par siARN, conduisant à un silençage transitoire de l'expression de *PRL* et ne permettant des observations qu'à court terme (2 jours de transfection). Il faudrait alors envisager une extinction plus stable de l'expression de *PRL*, par une stratégie de shARN par exemple, qui permettrait peut-être d'observer sur un temps plus long des phénotypes liés à la croissance cellulaire ou aux capacités migratoires.

Nous pouvons proposer d'autres stratégies expérimentales qui pourraient nous permettre d'affiner la description du phénotype de résistance accrue par les cellules *PRL*+. Nous pourrions tester, au-delà de l'actinomycine D, d'autres drogues induisant un stress ou une mort cellulaire, voire procéder à des tests de privations ou d'hypoxie sur les cellules CBPC. Nous pourrions également tenter un sauvetage (rescue) du phénotype de résistance, en faisant ré-exprimer par les cellules CBPC traitées par le siARN *PRL* une forme non-sécrétable de la prolactine, et en observant si l'on peut diminuer la mortalité induite par une drogue. Nous pouvons également proposer une stratégie *in vivo* : nous pourrions procéder à des tests de croissance tumorale dans des modèles de souris nude, en injectant dans les flancs des animaux des cellules *PRL*+ (H146 par exemple) et des cellules dans lesquelles l'expression du gène a été stablement supprimée (H146 shRNA *PRL*). Ces modèles nous permettraient de mesurer des paramètres tels que le temps de prise tumorale, la taille et l'invasivité des tumeurs ou encore la résistance de la tumeur à des traitements anti-cancéreux, selon que la tumeur exprime ou non le gène *PRL*.

Les analyses transcriptomiques, qu'elles portent sur nos lignées cellulaires ou sur les tumeurs, ne nous ont fourni pour l'instant qu'une vision globale des expressions géniques en absence ou en présence de l'expression de *PRL*. Ces analyses nous paraissent cohérentes, étant donné que les groupes de gènes identifiés dans deux lignées différentes, H146 et H524, présentent un recouvrement important. La comparaison avec les données issues de l'analyse transcriptomique des patients nous a également fourni un cœur de gènes dont le niveau d'expression semble être dépendant de l'expression du gène *PRL*. Malgré tout, ces approches bio-informatiques devront être confirmées expérimentalement, en choisissant quelques gènes cibles pertinents, et pourront à terme nous fournir des enseignements précieux pour décrypter les mécanismes à l'origine de l'acquisition d'une résistance accrue par les cellules cancéreuses pulmonaires *PRL*+

3.6. Hypothèses sur la protéine

Nous proposons que la production par les cellules cancéreuses pulmonaires d'un ARN messager *PRL* court, amputé de ses premiers exons en 5', puisse aboutir à la synthèse d'une protéine prolactine tronquée de son peptide signal, et retenue dans la cellule sans possibilité d'être sécrétée (**Figure 75**). Je n'ai pu fournir de preuve de l'existence de la protéine, ni de sa localisation intracellulaire. Nous avons proposé que l'anticorps utilisé ne soit pas capable de détecter la forme tronquée de la prolactine. Nous pouvons toutefois préciser qu'un certain nombre de protéines sont difficilement détectables dans les cellules cancéreuses pulmonaires. C'est le cas du facteur de transcription p53, présent en un très faible nombre d'exemplaires dans la cellule, qui est indétectable par immunohistochimie sur tumeur. Seule la forme mutée de la protéine, s'accumulant dans la cellule, devient détectable par les méthodes classiques (**Figure 22**) ([Brambilla et al., 1993](#)). Les protéines exprimées de manière aberrante ne sont pas souvent observées. C'est le cas par exemple de la protéine de fusion EML4-ALK qui est restée pendant un temps difficile à détecter ([Mino-Kenudson et al., 2010](#)).

Pour proposer malgré tout des indications sur le comportement qu'adopterait cette prolactine retenue dans les cellules cancéreuses pulmonaires *PRL*+, nous nous sommes appuyés sur une lignée modèle de cellules H1299 exprimant de manière stable la protéine de fusion GFP-PRL. Contrairement à sa protéine miroir PRL-GFP qui transite bien par l'appareil de Golgi et est sécrétée dans le milieu de culture, la protéine GFP-PRL est retenue dans la cellule, probablement du fait de l'incapacité de son peptide signal à assurer sa fonction d'orientation de

la protéine vers la voie de sécrétion du fait de la présence de la GFP. Nous avons tout d'abord observé que cette protéine de fusion était clivée, libérant un fragment carboxy-terminal de la prolactine d'environ 16 kDa. Ce n'est pas la première fois qu'un clivage de la prolactine est décrit : nous avons présenté le cas du clivage de l'hormone en vasoinhibines agissant sur les cellules endothéliales (**voir Contexte, paragraphe 1.3.1**). Néanmoins, ce clivage est encore mal compris, la protéase qui en est responsable est encore discutée et il semble avoir lieu dans le milieu extra-cellulaire. Dans notre cas, le clivage a lieu à l'intérieur de la cellule H1299, mais son existence réelle dans les lignées CBPC *PRL*+ reste à démontrer. L'observation qui nous a le plus intéressés est celle de l'acétylation subie par la GFP-*PRL* dans les cellules H1299.

L'acétylation des protéines en général, et pas uniquement celle des histones, joue un rôle important dans de nombreux processus biologiques ([Sadoul et al., 2011](#)) : dynamique du cytosquelette ([Zencheck et al., 2012](#)) ou métabolisme par exemple ([Xiong and Guan, 2012](#)). L'acétylation des lysines module la fonction de nombreuses protéines cytoplasmiques et nucléaires, dont la tubuline ou p53 par exemple. L'acétylation des protéines non-histones a été impliquée dans le cancer ([Arif et al., 2010](#) ; [Singh et al., 2010](#)), comme celle du suppresseur de tumeur RASSF5 ([Suryaraja et al., 2013](#)). L'acétylation de certaines protéines exprimées de manière ectopique a été observée dans notre équipe, par comparaison des gènes tissus-spécifiques activés dans les cancers ([Rousseaux, 2013](#)) et des données d'acétylome de notre collaborateur Minoru Yoshida (données non publiées).

La prolactine contient 12 résidus lysine qui pourrait potentiellement être acétylés. La cartographie des acétylations de la GFP-*PRL* pourrait être réalisée par spectrométrie de masse. L'enzyme responsable de son acétylation n'a pas été identifiée même si nous avons pu proposer p300 comme acétyltransférase. Si la protéine *PRL* est produite sous une forme tronquée dans les cellules cancéreuses pulmonaires *PRL*+, il serait particulièrement intéressant d'observer si elle est acétylée et déterminer le rôle de cette modification, sur sa localisation subcellulaire et sur sa fonction. En tous les cas, les résultats que nous avons obtenus au cours de cette thèse ont démontré pour la première fois l'expression du gène *PRL* au sein des cellules cancéreuses pulmonaires et nous permettent de proposer un rôle totalement inédit de cette expression ectopique dans la tumorigenèse pulmonaire. Alors que dans les cancers du sein et de la prostate, la prolactine sécrétée par les cellules tumorales agit localement via son récepteur dans une boucle autocrine/paracrine, nous avons affaire dans le poumon à un cas de figure totalement différent, où l'absence du récepteur et la production d'un ARNm *PRL* tronqué de ses séquences 5' nous conduisent à proposer un rôle *in cellulo* d'une prolactine non sécrétable pour promouvoir la résistance des cellules cancéreuses et plus globalement l'agressivité de la tumeur.

3.7. Utilisation de *PRL* comme biomarqueur ou cible thérapeutique ?

Nous avons démontré que le gène *PRL*, normalement réprimé dans le poumon sain, est exprimé de manière ectopique dans les tumeurs pulmonaires agressives. En conséquence, nous proposons que cette expression hors-contexte du gène *PRL* puisse être utilisée comme marqueur de mauvais pronostic. La détection de l'expression du gène *PRL* au sein de la tumeur, par PCR ou hybridation *in situ* sur échantillon biopsé, pourrait être utilisée par les cliniciens comme marqueur prédictif d'un pronostic sombre pour le patient. Une détection par analyse sur puce ne semble pas être une technique adaptée étant donné qu'elle fournit de faux-négatifs. De même, une détection dans les fluides (sérum notamment) ne semble pas être réalisable étant donné que la protéine prolactine n'est pas sécrétée. Quant aux méthodes de détection de la protéine, de type immunohistochimie sur coupes de tumeurs, elles nécessiteront d'abord de démontrer l'expression de la protéine à un niveau permettant la détection, et de trouver un anticorps capable de détecter efficacement la forme tronquée de cette protéine.

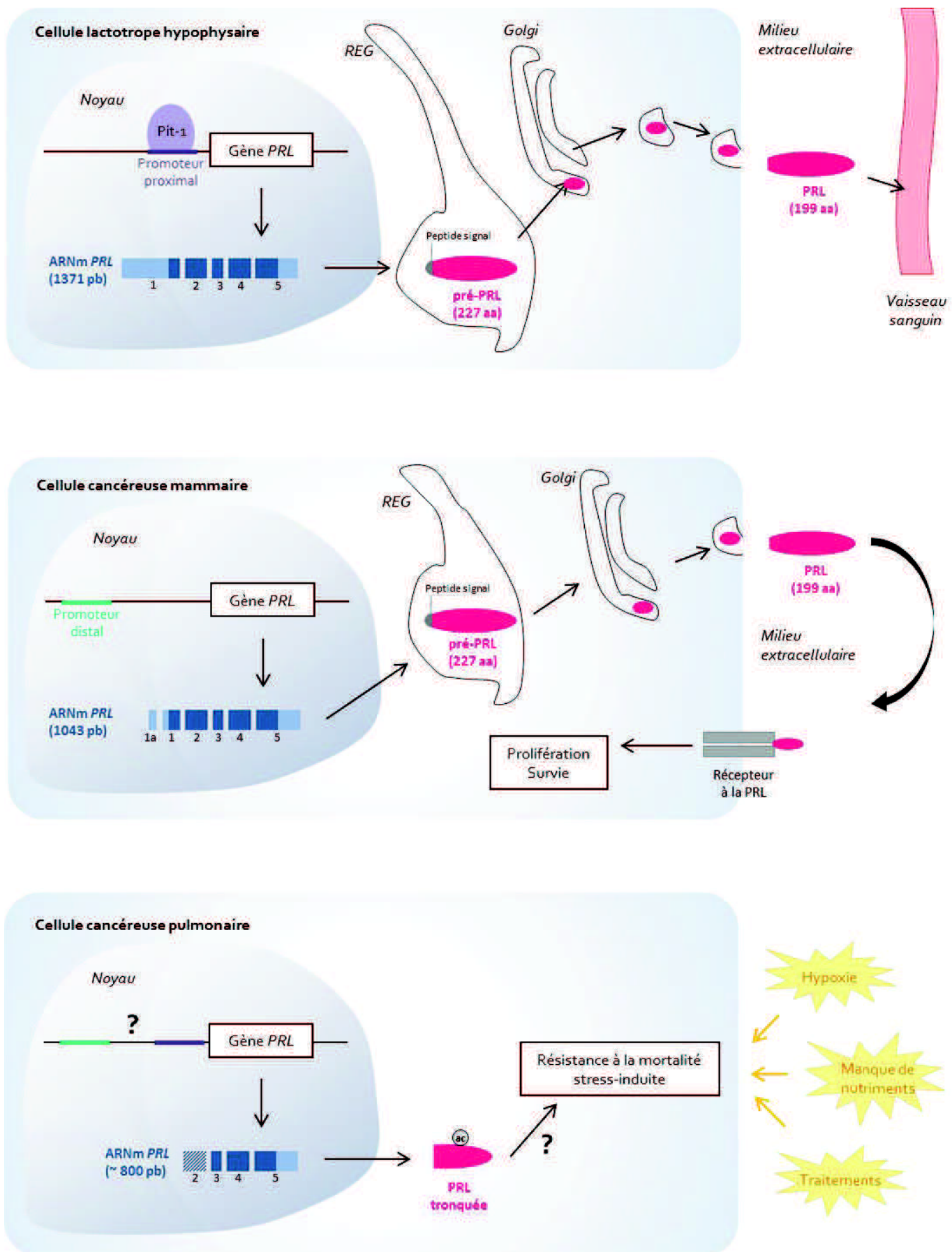


Figure 68 : Vers un mode d'action complètement inédit d'une prolactine tronquée non sécrétable agissant *in cellulo* pour conférer aux cellules cancéreuses pulmonaires une résistance accrue aux stress subis.

REG : réticulum endoplasmique granuleux ; ac : groupement acétyle

Puisque l'expression de *PRL* est à l'origine d'un phénotype agressif et accroît la résistance des cellules cancéreuses pulmonaires à la mort cellulaire, nous pouvons aussi proposer l'utilisation de l'expression hors-contexte de *PRL* comme cible thérapeutique. Néanmoins, nous sommes confrontés au même problème que celui soulevé dans les cancers du sein et de la prostate : celui de trouver un inhibiteur de l'expression extra-hypophysaire du gène, étant donné que les agonistes dopaminergiques utilisés pour réduire l'hyperprolactinémie systémique ne fonctionnent que sur les cellules lactotropes hypophysaires et pas sur les sites de production secondaires. De plus, nous sommes dans l'impossibilité d'utiliser des inhibiteurs du récepteur de la prolactine, actuellement en tests pour les cancers du sein et de la prostate, puisque le récepteur est absent des cellules cancéreuses pulmonaires. Toutefois, nous sommes en mesure de proposer un traitement réduisant de manière efficace l'expression du gène dans les cellules *PRL*⁺ : il s'agit d'un traitement par un inhibiteur de HDAC, la trichostatine A, qui a démontré une efficacité proche de celle des siRNA dans le silençage de l'expression *PRL*. Ce résultat suggère que l'expression de *PRL* possède un statut un peu particulier. En effet, les traitements par inhibiteurs de HDAC, en augmentant le niveau global d'acétylation des histones sur le génome, entraînent le plus souvent des augmentations d'expression de gènes. Le gène *PRL* quant à lui présente une réponse inverse, avec une réduction significative de la quantité d'ARN messager *PRL* produit après un traitement par la TSA. Les bases mécanistiques de cette baisse d'expression restent à élucider.

Plus généralement, il faut souligner que les inhibiteurs de HDAC sont suggérés depuis une quinzaine d'années comme des agents thérapeutiques très prometteurs dans la lutte contre les cancers ([Gridelli et al., 2008](#)). La FDA (U.S. Food and Drug Administration) a approuvé récemment la mise sur le marché de deux inhibiteurs de HDAC, le vorinostat (SAHA) et la romidepsin (FK228), pour le traitement de lymphomes à cellules T ([Duvic and Vu, 2007](#) ; [Marks and Breslow, 2007](#)) et plus d'une vingtaine d'inhibiteurs de HDAC différents sont en tests cliniques pour des pathologies néoplasiques variées (tumeurs solides et lymphomes). Dans les cas de cancers du poumon non à petites cellules, les inhibiteurs de HDAC pivanex, CI-994, vorinostat et LBH589 notamment, sont en évaluation clinique ([Witta, 2012](#)). On a démontré l'activité antitumorale des inhibiteurs de HDAC dans les CBNPC ([Mukhopadhyay et al., 2006](#)), notamment par induction de l'apoptose ([Choi, 2005](#)). Il semble que leur effet soit permis par la réexpression de gènes suppresseurs de tumeur précédemment réprimés ([Brehm et al., 1998](#)), mais également par des effets indépendants de la transcription. Les inhibiteurs de HDAC semblent cibler davantage les cellules cancéreuses que les cellules saines, même si les raisons de cette sélectivité sont encore débattues ([Lee et al., 2012](#)).

Au-delà de l'effet de la TSA sur l'expression du gène *PRL*, nous avons démontré que les cellules pulmonaires cancéreuses *PRL*⁺ étaient particulièrement sensibles à un traitement par cet inhibiteur de HDAC, la TSA entraînant par exemple l'apoptose de plus de 60% des cellules H146. Ceci suggère que les inhibiteurs de HDAC pourraient être utilisés en thérapeutique pour induire de manière sélective la mort des cellules cancéreuses *PRL*⁺. On pourrait confirmer l'intérêt d'un tel traitement dans un modèle *in vivo* de xénogreffe en souris nude, en testant l'effet d'un traitement aux inhibiteurs de HDAC (seul ou en combinaison avec un autre traitement anti-cancéreux) sur la croissance de tumeurs *PRL*⁺. A terme, l'expression hors-contexte du gène *PRL* pourrait donc être utilisée à la fois comme biomarqueur de pronostic sombre et comme biomarqueur de réponse à un traitement par inhibiteur de HDAC.



**Expression aberrante du effecteur épigénétique testiculaire BRDT
dans les tumeurs pulmonaires :**

Un marqueur de réponse au traitement par les inhibiteurs de HDAC

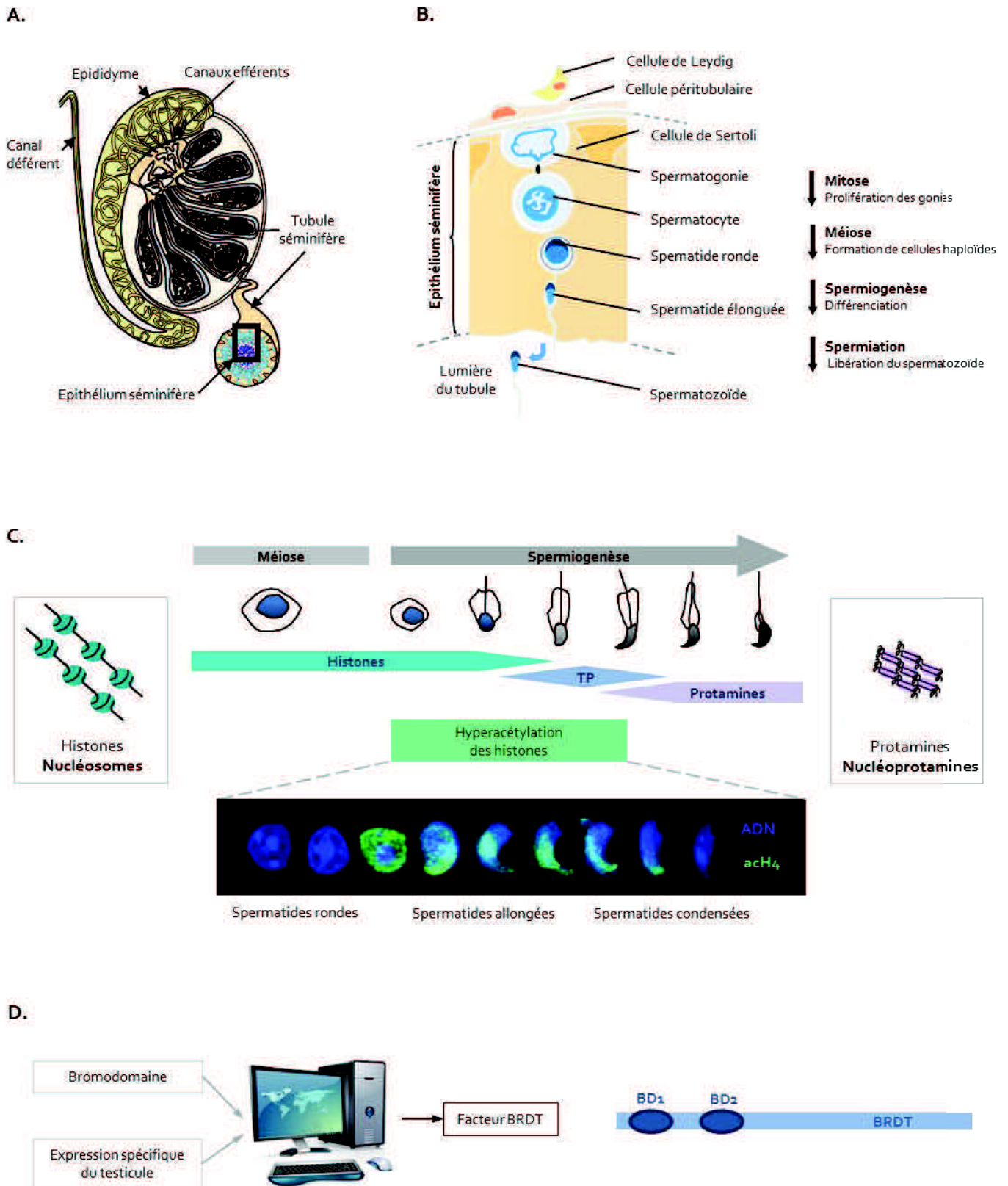


Figure 69 : BRDT, un facteur à bromodomaine spécifique de la lignée germinale mâle

A. Anatomie du testicule. Les tubules séminifères contiennent les cellules germinales soutenues par les cellules de Sertoli.

B. La différenciation des cellules germinales a lieu de manière centripète. Les gonies prolifèrent puis la méiose donne naissance à des spermatocytes haploïdes. Les spermatides subissent une différenciation drastique jusqu'à libération du spermatozoïde.

C. Lors des phases post-méiotiques de la spermatogenèse, les histones sont enlevées et remplacées par des protéines de transition (TP) puis des protamines. Cet échange est précédé par une vague globale d'acétylation des histones dans les spermatides en allongement, observable en immunofluorescence grâce à un anticorps anti-histone H₄ acétylée.

D. Une analyse *in silico* permet d'identifier le facteur BRDT, protéine spécifique du testicule et qui possède deux bromodomaines (BD), comme candidat pouvant participer à la ré-organisation de l'épigénome mâle.

1. Contexte de l'étude

La spermatogenèse est un processus physiologique qui aboutit à la formation de gamètes haploïdes, les spermatozoïdes, assurant le transport du génome mâle dans les voies génitales femelles jusqu'à l'ovaire. Ce processus complexe, prenant place dans les tubules séminifères du testicule où a lieu la différenciation des cellules de la lignée germinale mâle (**Figure 69A et B**), implique des changements épigénétiques drastiques ([Moriniere et al., 2009](#)). Lors des étapes post-méiotiques de la spermatogenèse a lieu une reprogrammation massive du génome mâle, l'évènement le plus marquant étant l'éviction de 90% des histones et l'emballage du génome par de petites protéines basiques, les protamines (**Figure 69C**). Cette transition d'une organisation nucléosomale de la chromatine vers une organisation en nucléoprotamines permet une compaction drastique du génome dans la tête du spermatozoïde. Néanmoins, les mécanismes physiologiques à l'origine de cette réorganisation restent assez mal connus.

Les travaux de notre équipe ont mis en évidence l'existence d'une vague d'acétylation des histones dans les spermatides en élongation, précédant l'enlèvement des histones ([Govin et al., 2004](#)) (**Figure 69C**). L'hypothèse a alors été émise que cette hyperacétylation globale puisse conditionner l'éviction des histones en recrutant des effecteurs épigénétiques sur les résidus acétylés. Les bromodomaines avaient alors été décrits comme des domaines protéiques capables d'interagir spécifiquement avec les histones acétylées ([Dhalluin et al., 1999](#); [Jacobson et al., 2000](#)) (**Figure 6C**). Dans ce contexte, l'équipe a lancé une analyse *in silico* à la recherche d'éventuels facteurs à bromodomaine dont l'expression serait spécifique du testicule, et plus précisément des cellules de la lignée germinale mâle subissant la vague d'acétylation des histones. Cette analyse a abouti à l'identification d'un candidat potentiel, le facteur BRDT (pour BromoDomain Testis-specific) (**Figure 69D**), une protéine spécifique des cellules post-méiotiques (**Figure 11B**) possédant deux bromodomaines ([Jones et al., 1997](#)).

Dans les années qui ont suivi, les travaux au sein de l'équipe ont permis de mettre en lumière le rôle majeur assuré par BRDT dans la reprogrammation du génome mâle. On a démontré que le premier bromodomaine de BRDT (BD1) est capable de cibler les nucléosomes hyperacétylés ([Govin et al., 2004](#)). Il interagit spécifiquement avec l'histone H4 acétylée sur ses résidus lysines K5 et K8 ([Moriniere et al., 2009](#)) (**Figure 70A**), ce qui fait de ce facteur le premier « lecteur épigénétique » jamais décrit capable de lier combinatoirement des modifications post-traductionnelles sur plusieurs résidus d'histone. L'acétylation de H4 ayant lieu des résidus carboxy-terminaux vers les résidus de l'extrémité amino-terminale (**Figure 70A**), les acétylations K5 et K8 ne sont observées qu'en présence conjointe des acétylations de K12 et K16 : BRDT interagit donc spécifiquement avec la forme hyperacétylée de H4. Les souris exprimant la protéine BRDT privée de son premier bromodomaine sont stériles ([Shang et al., 2007](#)). La spermatogenèse, défectueuse, est caractérisée par une persistance anormale d'histones dans les cellules germinales les plus différenciées, une accumulation cytoplasmique des protamines qui ne sont pas correctement intégrées au génome et un défaut de compaction des spermatides ([Gaucher et al., 2012](#)) (**Figure 70B**). Enfin, une mutation homozygote du gène *BRDT* a été identifiée chez des hommes infertiles ([Rousseaux and Khochbin, 2009](#)). Ces résultats dans leur ensemble illustrent la capacité de BRDT à se lier aux nucléosomes ayant subi la vague d'hyperacétylation et à conditionner le bon enlèvement des histones nécessaire à la production des gamètes (**Figure 70B**).

Dans le même temps, on a observé que le gène *BRDT*, dont l'expression est normalement restreinte aux spermatides ([Shang et al., 2007](#)), est ré-exprimé de manière aberrante dans 20% des tumeurs (**Figure 70C**), faisant de BRDT un nouveau facteur CT, le facteur CTg ([Scanlan et al., 2002b](#)). En particulier, la présence de l'ARN codant BRDT a été rapportée dans 12 cas (sur 47 étudiés) de cancers du poumon non à petites cellules, et dans un cas de carcinome à cellules squameuses. Etant donné le rôle de BRDT dans la réorganisation de l'épigénome mâle, la question s'est posée des conséquences de son expression dans des cellules somatiques

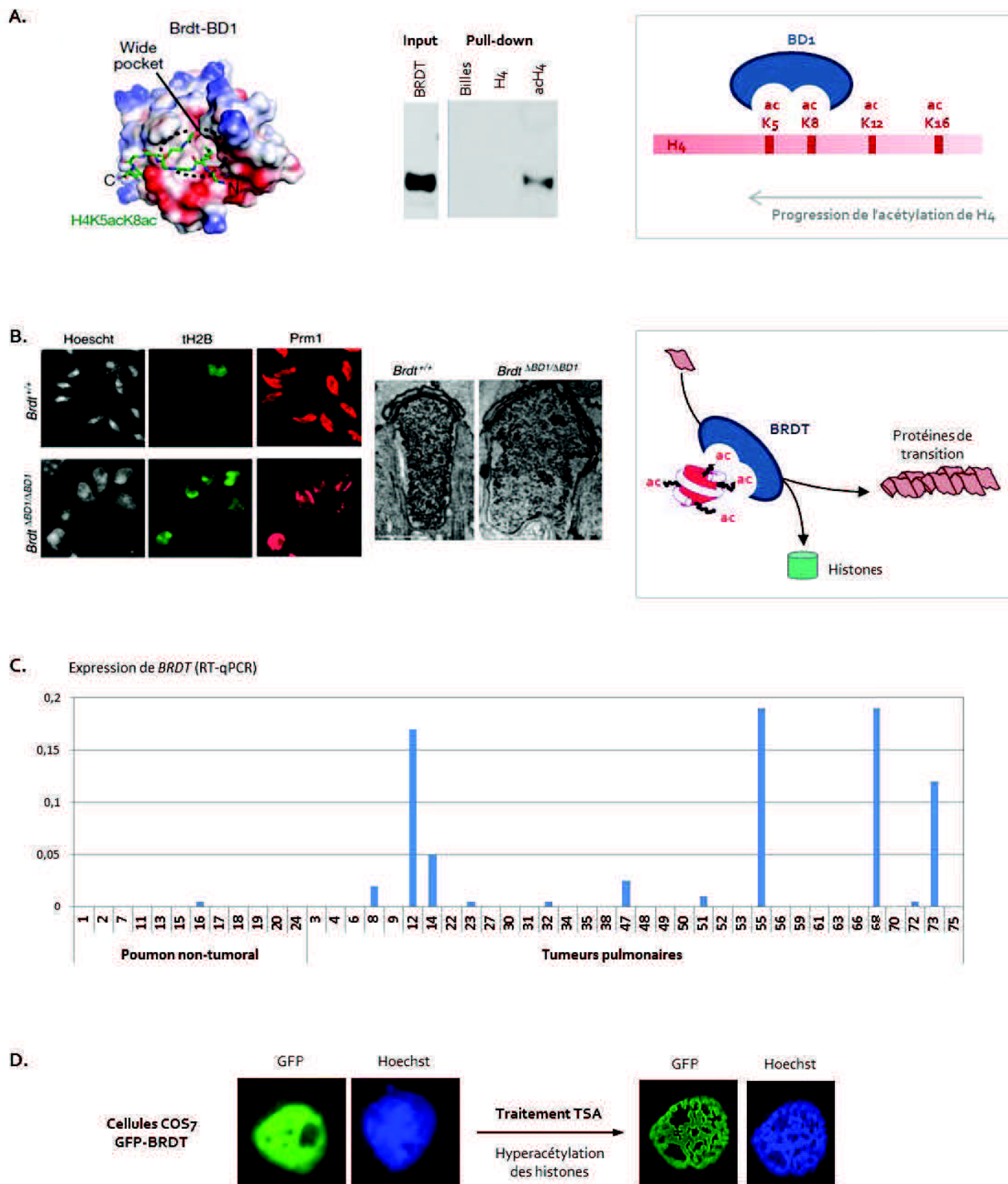


Figure 70 : BRDT est impliqué dans le remodelage de la chromatine acétylée dans les cellules germinales mais également dans les cellules somatiques lorsqu'exprimé ectopiquement

A. (Panneau de gauche) Structure du BD1 interagissant avec la queue N-terminale de H₄ acétylée sur les résidus K5 et K8. Tiré de (Moriniere et al., 2009). (Panneau du milieu) Liaison de BRDT à la forme acétylée de l'histone H₄ démontrée en pull-down. Tiré de (Gaucher et al., 2012). (Panneau de droite) L'acétylation de K5 et K8 signe l'hyperacétylation de l'histone H₄.

B. (Panneau de gauche) Les souris exprimant BRDT délété de son premier bromodomaine (Δ BD1) ont des spermatides anormalement compactées, avec des histones qui persistent et un défaut d'incorporation des protamines. Tiré de (Gaucher et al., 2012). (Panneau de droite) BRDT est donc impliqué dans l'enlèvement des histones et leur remplacement.

C. L'expression de *BRDT* est détectée dans certaines tumeurs pulmonaires de la cohorte (Elisabeth Brambilla, Sylvie Gazzeri).

D. Des cellules Cos7 sont transfectées transioirement avec un vecteur codant pour la protéine BRDT qui se localise dans le noyau. L'hyperacétylation induite par un traitement à la trichostatine A cause une compaction anormale de la chromatine.

cancéreuses : on peut imaginer que son expression hors-contexte puisse contribuer au bouleversement de l'épigénome dans ces cellules, conduisant éventuellement à la transformation tumorale ([Rousseaux and Khochbin, 2009](#)). Des travaux amorcés dans l'équipe ont montré que l'expression aberrante de BRDT dans des cellules somatiques, associée à une hyperacétylation des histones obtenue par un traitement avec un inhibiteur des histones déacétylases (HDACi), la trichostatine A (TSA), aboutit à un effondrement de la structure chromatinienne, caractérisée par une compaction anormale du génome observable après marquage au Hoechst en microscopie à épifluorescence conduisant à la mort cellulaire (**Figure 70D**) ([Pivot-Pajot et al., 2003](#)).

C'est dans ce contexte que nous nous sommes demandé si cette capacité de BRDT, exprimé ectopiquement dans des cellules somatiques, à causer une compaction dramatique et létale de la chromatine hyperacétylée, ne pourrait pas être utilisée contre les cellules cancéreuses exprimant BRDT ([Reynoird et al., 2010a](#)). L'objectif du travail préliminaire que j'ai mené a été de déterminer si une hyperacétylation des histones induite par la trichostatine A dans les cellules cancéreuses pulmonaires exprimant BRDT de manière endogène pourrait conduire à une apoptose massive. Cette hypothèse, si elle était vérifiée, pourrait nous amener à proposer l'activation illégitime de BRDT au sein d'une tumeur broncho-pulmonaire comme marqueur de réponse à un traitement par HDACi.

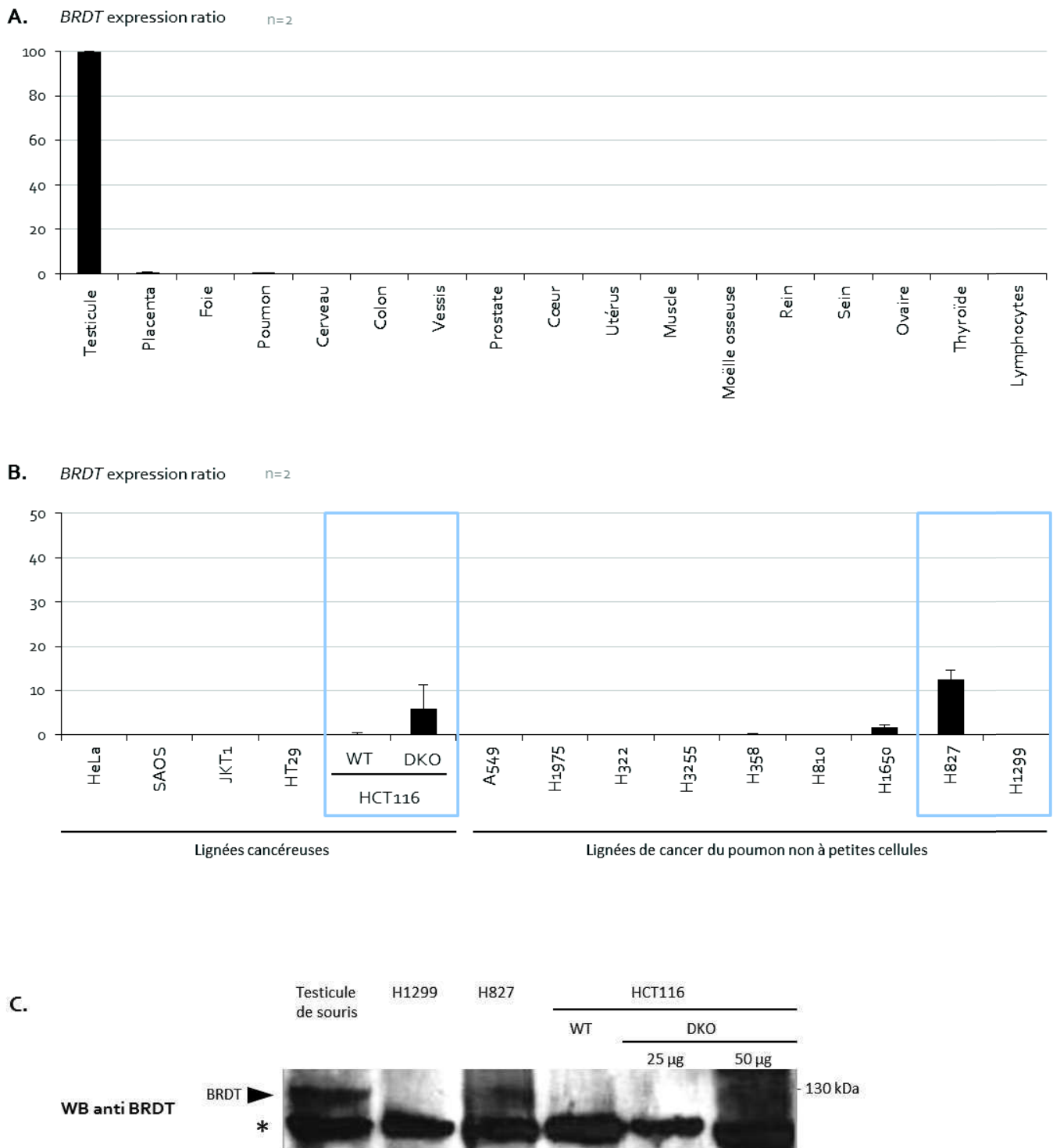


Figure 71 : BRDT est un facteur spécifique du testicule dont l'expression ectopique est détectée dans des lignées cancéreuses pulmonaires

A. Expression du gène *BRDT* dans un panel de tissus humains sains. Ces résultats ont été obtenus par deux expériences indépendantes de RT-qPCR avec normalisation par quatre gènes contrôle (*Actine*, *U6*, *RELA* et *AUP1*). Les barres d'erreur représentent les écarts-type. Les primers utilisés sont spécifiques des deux variants d'ARN *BRDT*. 100 % = Ratio d'expression dans le testicule.

B. Expression du gène *BRDT* dans un panel de lignées cancéreuses humaines. Ces résultats ont été obtenus par deux expériences indépendantes de RT-qPCR avec normalisation par quatre gènes contrôle (*Actine*, *U6*, *RELA* et *AUP1*). Les barres d'erreur représentent les écarts-type. Les primers utilisés sont spécifiques des deux variants d'ARN *BRDT*. 100 % = Ratio d'expression dans le testicule.

C. Détection de la protéine BRDT grâce à un anticorps anti-BRDT dans des extraits protéiques totaux de lignées de cancer du poumon. Le contrôle positif est fourni par un extrait total de testicule de souris. *: bande aspécifique.

2. Résultats

2.1. La protéine BRDT spécifique de la lignée germinale mâle est produite ectopiquement dans une lignée de cancer pulmonaire

Dans un premier temps, j'ai souhaité confirmer le patron d'expression du gène *BRDT*, notamment dans le but d'identifier une lignée cellulaire exprimant de manière aberrante le gène et la protéine correspondante (**Figure 71**). Dans un panel de tissus humains sains, il n'y a que dans le testicule que l'on détecte l'expression de *BRDT* en RT-qPCR. On remarque plus particulièrement que ce gène est réprimé dans le poumon sain (**Figure 71A**). L'analyse de l'expression de *BRDT* au sein de plusieurs lignées cellulaires cancéreuses montre une expression dans deux lignées de cancers du poumon non à petites cellules (**Figure 71B**). L'expression est relativement faible dans la lignée H1650, mais plus marquée dans la lignée H827 où le niveau d'expression dépasse 10% de l'expression testiculaire. On remarque également une réactivation du gène *BRDT* dans la lignée HCT116 DKO chez laquelle les activités DNMT1 et 3B ont été supprimées, causant une déméthylation globale du génome, alors qu'aucune expression n'était détectée dans la lignée parentale HCT116 WT. Ce résultat suggère que l'activation ectopique du gène *BRDT* dans les cellules cancéreuses somatiques puisse être due à une déméthylation aberrante de son promoteur. Au niveau protéique (**Figure 71C**), on observe la présence de la protéine murine dans un extrait de testicule aux alentours de sa taille attendue (130 kDa). La même bande est détectée dans la lignée H827, et plus difficilement dans la lignée HCT116 DKO. A l'inverse, la protéine BRDT n'est pas observée dans l'extrait protéique de cellules H1299. On utilisera alors dans la suite de ce projet la lignée H827 comme lignée de cancer du poumon BRDT-positive, et la lignée H1299 comme lignée BRDT-négative.

2.2. La présence ectopique de BRDT induit une plus grande sensibilité des cellules cancéreuses à un traitement par la TSA

Une compaction drastique de la chromatine, causant une mort cellulaire, a été observée suite à un traitement par la TSA dans des cellules COS exprimant transitoirement la protéine BRDT. J'ai voulu savoir si cet effet de la TSA était observable dans les cellules cancéreuses pulmonaires exprimant de manière endogène la protéine BRDT. En particulier, je me suis demandé si l'expression de BRDT dans les cellules H827 les rendrait plus sensibles à l'apoptose induite par la TSA. J'ai mesuré l'effet d'un traitement par la TSA ou par l'actinomycine D sur la mortalité des cellules en cytométrie de flux (**Figure 72B**). Les cellules des deux lignées H827 (BRDT-positive) et H1299 (BRDT-négative) présentent une mortalité basale en condition contrôle comparable (environ 2,5% de cellules apoptotiques). Nous constatons que les cellules H827 BRDT-positives sont plus sensibles à l'apoptose induite par la TSA et par l'actinomycine D que les cellules H1299, avec des proportions de cellules apoptotiques de 22 contre 4% pour la TSA et 49 contre 30% pour l'actinomycine D. De manière cohérente, la forme clivée de PARP, marqueur d'apoptose, apparaît plus intense en Western blot dans les cellules H827 que dans les cellules H1299 suite aux traitements (**Figure 72C**).

Pour savoir si BRDT est impliqué dans cette mortalité induite par les drogues, j'ai réalisé un silençage de l'expression de *BRDT* par la technique d'interférence ARN (**Figure 72A**). La transfection du siARN BRDT cause une diminution de 85% de l'expression du gène en qPCR et la présence de la protéine n'est plus détectable en Western blot, sans modification de la morphologie cellulaire. On peut également noter qu'un traitement par la TSA ne modifie pas le niveau d'expression de *BRDT*. La quantification des cellules apoptotiques en cytométrie (**Figure 72B**) nous révèle que la mortalité basale des cellules H827 n'est pas affectée par le silençage du gène *BRDT* et reste stable aux alentours de 3%. Par contre, les cellules H827 transfectées par le siARN BRDT subissent moins d'apoptose que les cellules H827 transfectées avec le siARN contrôle après traitement à la

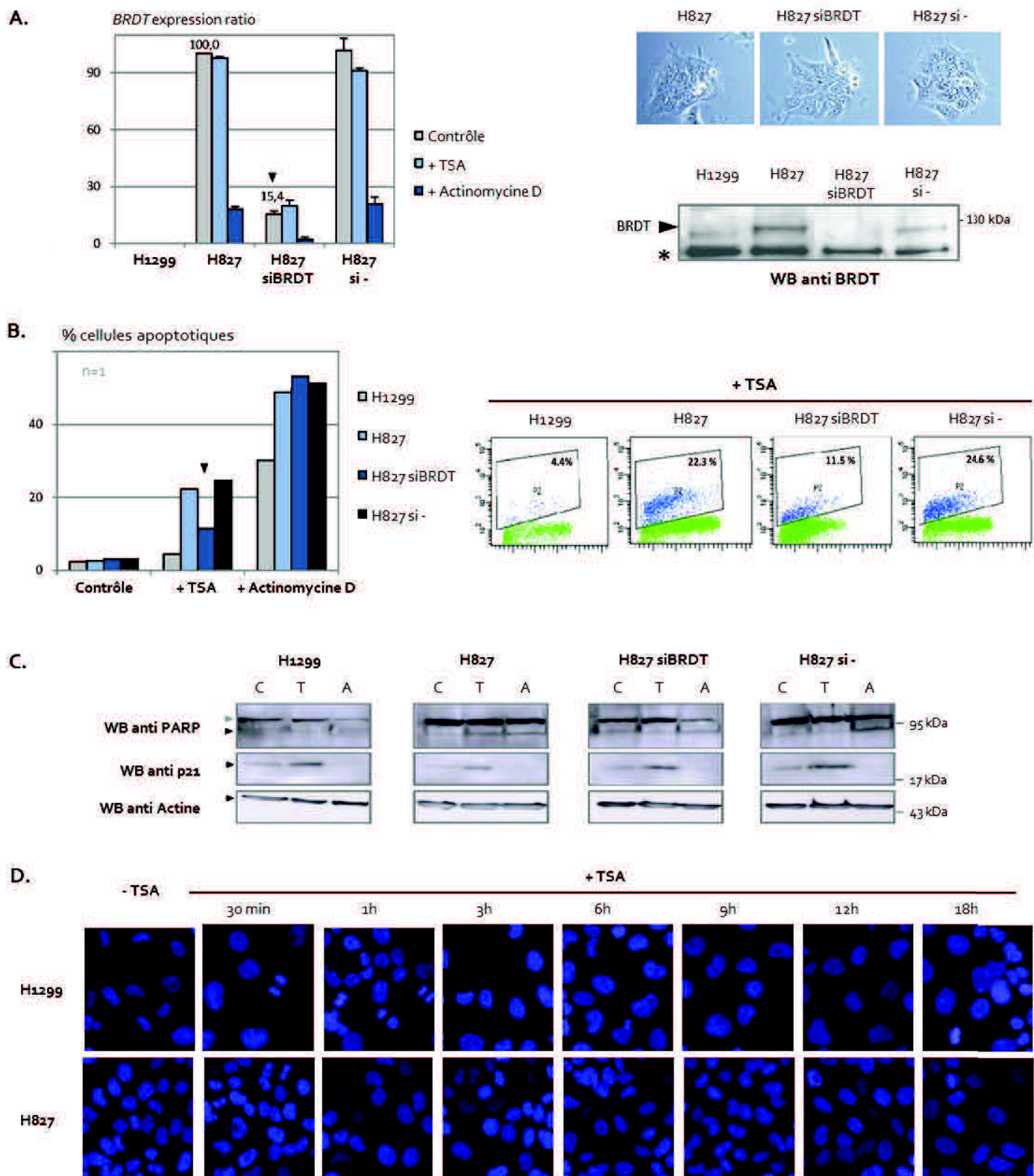


Figure 72 : BRDT est impliqué dans la mortalité induite par la TSA des cellules cancéreuses pulmonaires H827

A. (Panneau de gauche) Expression du gène *BRDT* dans les lignées cancéreuses pulmonaires H1299 et H827, après éventuelle transfection du siARN *BRDT*, et traitement durant 12h par la TSA (100 ng/mL) ou par l'actinomycine D (250 nM), obtenue par 2 expériences indépendantes de RT-qPCR avec normalisation par 2 gènes contrôle (*Actine* et *U6*). 100 % = Ratio d'expression dans la lignée H827. si - : contrôle négatif grâce à un siARN scramble. (Panneau de droite en haut) Effet du siARN sur la morphologie des cellules H827. (Panneau de droite en bas) Détection de la protéine *BRDT* en Western blot dans les 4 lignées non traitées.

B. Quantification de l'apoptose par cytométrie en flux grâce à un anticorps anti-caspase 3 activée dans les différentes populations cellulaires H1299 et H827, transfectées ou non avec le ARN *BRDT*, traitées ou non par la TSA (100 ng/mL) ou par l'actinomycine D (250 nM). Les données de cytométrie après traitement à la TSA sont présentées.

C. Détection des protéines *BRDT*, *PARP* entière (116 kDa) ou clivée, marqueur d'apoptose (89 kDa), *p21* (activée par un traitement à la TSA) et *Actine* (contrôle de charge) en Western blot. C = contrôle, T = TSA, A = Actinomycine D.

D. Observation par microscopie à fluorescence des cellules H1299 et H827 après marquage de l'ADN par le Hoechst lors d'une cinétique de traitement par la TSA.

TSA (12 contre 22%), un effet qui n'est pas observé suite au traitement à l'actinomycine D. Etant donné la diminution de la mortalité induite par la TSA observée lorsque l'expression de *BRDT* est réduite, nous concluons que *BRDT* contribue à la sensibilité des cellules H827 à l'apoptose causée par un traitement à la TSA. A l'inverse, *BRDT* ne semble pas impliqué dans l'apoptose induite par l'actinomycine D.

Cette mortalité induite par la TSA pourrait être liée à une compaction anormale de la chromatine hyperacétylée médiée par la protéine *BRDT* exprimée hors de son contexte. Néanmoins, nous n'observons aucune figure caractéristique de l'effondrement de la structure chromatinienne dans les cellules H827 *BRDT*-positives marquées par le Hoechst, et ce quelle que soit la durée de traitement par la TSA (**Figure 72D**). Nous émettons alors l'hypothèse que le faible niveau d'expression de cette protéine *BRDT* dans les cellules cancéreuses pulmonaires, bien en deçà du niveau d'expression obtenu par la transfection transitoire de cellules COS (**Figure 70D**), ne soit pas suffisante pour induire un remodelage drastique de la chromatine, tout en étant capable d'induire la mort des cellules par des effets plus discrets.

2.3. Etablissement d'une lignée cancéreuse pulmonaire modèle exprimant *BRDT* de manière stable

Lors de ces travaux préliminaires, j'ai mis au point un outil qui pourra être utile pour des analyses plus poussées du rôle de *BRDT* dans les cellules cancéreuses pulmonaires. Il s'agit de lignées cellulaires stables, établies à partir de la lignée H1299 *BRDT*-négative de cancer du poumon non à petites cellules, exprimant des formes de *BRDT* fusionnées à la GFP (**Figure 73**). J'ai travaillé dans la lignée H1299 qui n'exprime pas de *BRDT* endogène (**Figure 73B et C**), avec la forme entière de *BRDT* (appelée FL pour « forme longue ») et avec une forme de *BRDT* amputée de son domaine C-terminal mais contenant les deux bromodomains (appelée ΔC). Les lignées stables ont été obtenues par transfection des vecteurs, sélection des cellules ayant intégré le transgène par un antibiotique et tri de cellules fluorescentes en cytométrie de flux. Une lignée stable exprimant uniquement la GFP a été créée conjointement comme contrôle négatif. Les trois protéines peuvent être visualisées par Western blot avec un anticorps reconnaissant la GFP (**Figure 73A**) : la GFP est détectée à une taille apparente de 30 kDa, *BRDT* ΔC -GFP apparaît sous la forme d'un doublet autour de 95 kDa et *BRDT* FL-GFP est observée au-delà de 130 kDa. Les niveaux d'expression sont relativement faibles dans ces lignées stables ayant intégré le transgène par rapport à la surexpression massive obtenue en transfection transitoire.

On observe en cytométrie que 90% des cellules des lignées stables expriment le transgène. Alors que la GFP adopte une localisation cytoplasmique et nucléaire, les formes courtes et longues de *BRDT* adoptent une localisation nucléaire exclusivement (**Figure 73B**). Le marquage nucléaire est homogène dans les cellules H1299 *BRDT* ΔC -GFP en interphase, mais on observe un marquage des chromosomes mitotiques dans les cellules en division. Cette observation, réalisée uniquement pour *BRDT* ΔC -GFP et pas pour la GFP seule dans la lignée contrôle, est en accord avec la capacité déjà décrite des protéines de la famille BET (Bromo and Extra Terminal proteins), dont fait partie *BRDT*, à s'associer aux chromosomes mitotiques ([Dey et al., 2003](#)). La morphologie et le cycle cellulaire n'apparaissent pas modifiés par l'expression ectopique des protéines de fusion (données non présentées).

J'ai poussé un peu plus en avant la caractérisation de la forme *BRDT* ΔC -GFP. J'ai testé l'hypothèse selon laquelle la présence d'un doublet de bandes en Western blot puisse être due à une modification post-traductionnelle de la protéine. Je montre que la bande de poids moléculaire le plus élevé disparaît au profit de la bande la plus basse suite à un traitement par une phosphatase (**Figure 73C**), ce qui suggère que la protéine *BRDT* ΔC -GFP puisse être phosphorylée. Je me suis ensuite demandé si la protéine *BRDT* ΔC -GFP exprimée de manière aberrante dans des cellules cancéreuses pulmonaires conservait sa capacité de lier l'histone H4 acétylée. Par une expérience de peptide pull-down, je montre que les deux bandes correspondant à *BRDT* ΔC -

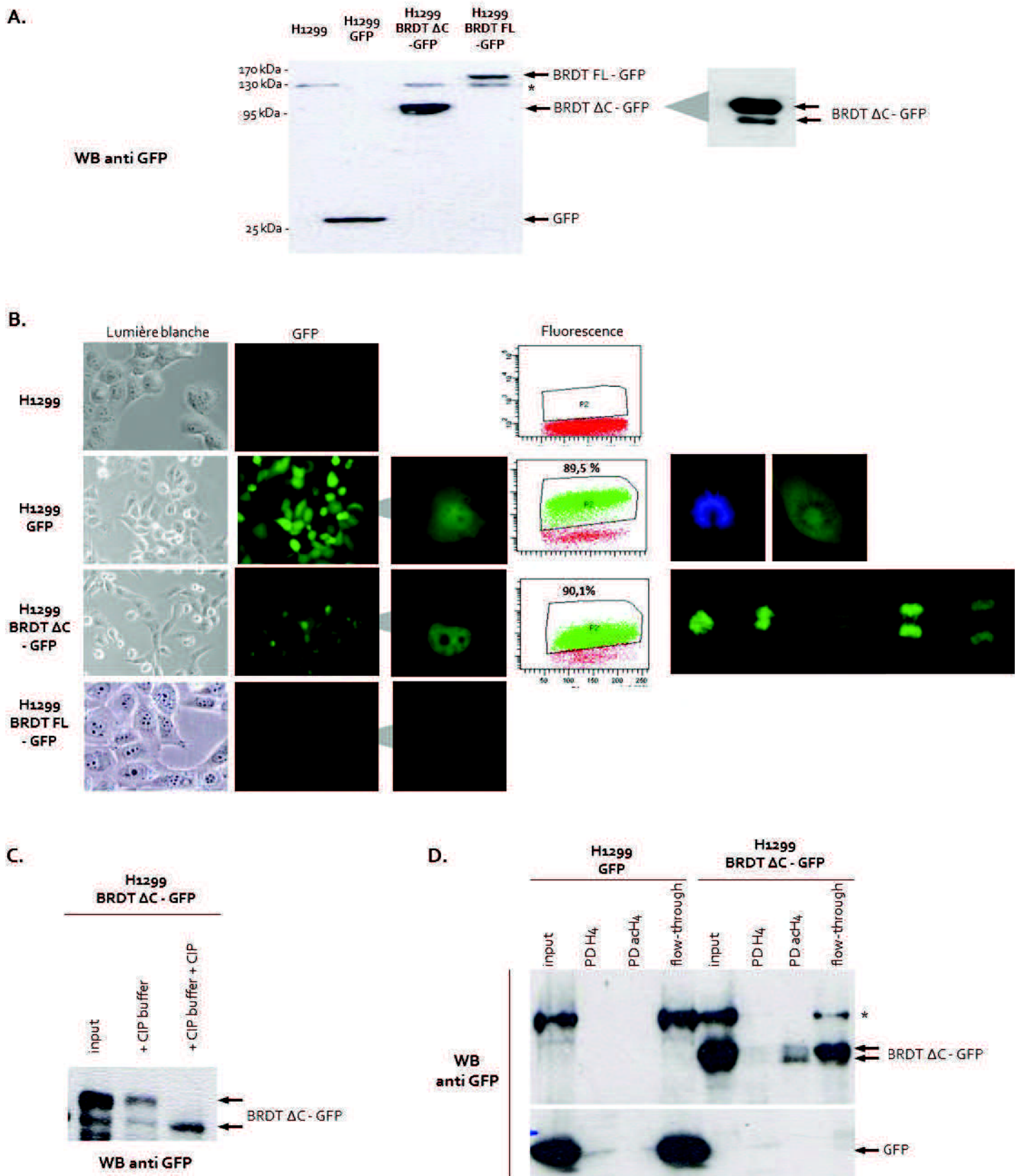


Figure 73 : Caractérisation des lignées stables H1299 exprimant des formes de BRDT fusionnées à la GFP

A. Détection en Western blot avec un anticorps anti-GFP des formes GFP, BRDT Δ C-GFP et BRDT FL-GFP exprimées de manière stable dans les lignées cellulaires établies.

B. Observation en microscopie à épifluorescence des cellules des trois lignées stables (et de la lignée mère H1299) et quantification de la proportion de cellules fluorescentes par cytométrie de flux.

C. La forme BRDT Δ C-GFP apparaît sous la forme d'un doublet en western blot. La bande de plus haut poids moléculaire n'est plus observable après une incubation du lysat cellulaire avec la phosphatase CIP.

D. Peptide-pull down réalisé sur des lysats de la lignée contrôle H1299 et de la lignée stable H1299 BRDT Δ C-GFP. Le doublet BRDT Δ C-GFP interagit avec un peptide H₄ acétylé. Aucune interaction n'est détectée avec le peptide H₄ non acétylé.

GFP (formes phosphorylée et non-phosphorylée) sont capables d'interagir avec un peptide H₄ tétra-acétylé (**Figure 73D**), mais pas avec le peptide non-acétylé. Se pose alors la question des conséquences que pourrait avoir l'hyperacétylation des histones induite par un traitement à la TSA sur les cellules de ces lignées stables.

2.4. Un traitement par la TSA de la lignée modèle H1299 GFP-BRDT induit un effondrement de la chromatine hyperacétylée

Après 12h de traitement de la lignée stable H1299 BRDT Δ C-GFP par l'inhibiteur des déacétylases TSA, on observe une augmentation de l'intensité de la fluorescence des cellules en microscopie et en cytométrie de flux, pouvant signer une augmentation de la synthèse ou une accumulation de la protéine de fusion, effet qui n'est pas observé dans la lignée contrôle H1299 GFP (**Figure 74A**). Cette accumulation de BRDT Δ C-GFP est également observée à l'échelle protéique en western-blot où le doublet apparaît beaucoup plus abondant dans la condition TSA 12h (**Figure 74C**).

Pour préciser le comportement de la protéine BRDT Δ C-GFP au cours de l'hyperacétylation induite par la TSA, l'analyse en western blot a été réalisée sur des extraits protéiques fractionnés (**Figure 74C**). Un protocole de fractionnement sub-cellulaire permet de séparer la fraction cytosoluble des protéines libérées par une simple permabilisation membranaire d'une part, et les protéines nucléaires insolubles (solubilisées uniquement par sonication dans un tampon dénaturant hypertonique) comme l'histone H₃. Alors que dans les conditions contrôle la protéine BRDT Δ C-GFP est obtenue dans la fraction soluble, elle est observée dans la fraction nucléaire insoluble dès 6 heures de traitement par la TSA. Ce résultat suggère que l'hyperacétylation des histones permet l'association de BRDT Δ C-GFP avec la chromatine, en accord avec la capacité de BRDT de lier l'histone H₄ acétylée (**Figure 74D**).

Le résultat le plus marquant est l'observation de figures de compaction de la chromatine, obtenues dès 6 heures de traitement à la TSA (**Figure 74B**). Le marquage par le Hoechst met en lumière un effondrement de la chromatine dans les cellules H1299 BRDT Δ C-GFP, chromatine qui prend un aspect grumuleux après le traitement par l'inhibiteur de HDAC. Ainsi, on reproduit en lignée pulmonaire stable le phénotype de compaction qui avait précédemment été observé en transfection transitoire dans les cellules COS7. Pour tenter d'obtenir des pistes sur les causes et les conséquences moléculaires de cette effondrement chromatinien, j'ai étudié les niveaux d'expression de quelques protéines d'intérêt en western blot (**Figure 74B**) dans les lignées H1299 GFP et H1299 BRDT Δ C-GFP. Alors que l'actine est utilisée comme témoin de charge, l'induction de p21 est un contrôle de l'efficacité du traitement TSA. Notre analyse grâce à un anticorps spécifique des formes acétylées d'H₄ montre une différence de comportement de cette modification d'histone dans la lignée H1299 BRDT Δ C-GFP comparée à la lignée contrôle H1299 GFP. En effet, alors que l'acétylation de H₄ n'est observable dans la lignée contrôle qu'à partir de 12 heures de traitement par la TSA, elle est observable dès 6 heures dans la lignée exprimant BRDT Δ C-GFP. Ce résultat pourrait suggérer un rôle de BRDT dans la maintenance de l'acétylation de H₄, la fixation de BRDT à ce résidu modifié pouvant le stabiliser. Quant au signal de l'histone H₃, il présente une baisse marquée dans la lignée H1299 BRDT Δ C-GFP après 6 heures de traitement par l'HDACi. Pourtant, cette baisse de signal ne semble pas être due à une baisse de la quantité globale d'histone H₃ dans la cellule, étant donné que la protéine H₃ est visible par coloration au bleu de Coomassie. On peut émettre l'hypothèse que des modifications post-traductionnelles de H₃ limitent sa reconnaissance par l'anticorps anti-H₃.

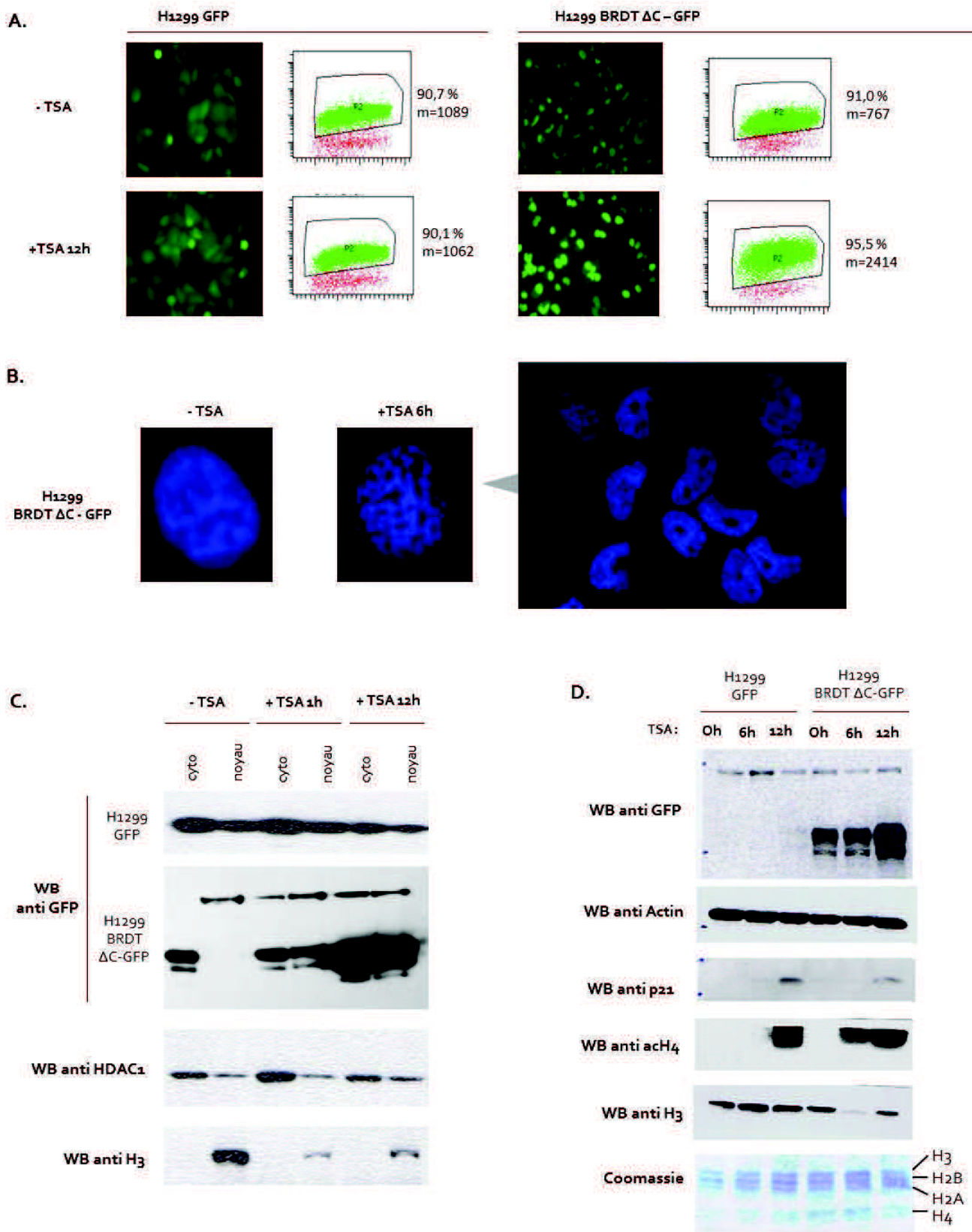


Figure 74 : La TSA induit l'augmentation d'expression de BRDT et son accrochage à la chromatine, associés à une compaction anormale de la chromatine dans le clone H1299 BRDT ΔC-GFP

A. Les cellules des lignées stables H1299 GFP et H1299 BRDT ΔC-GFP sont soumises à un traitement par la TSA et observées en microscopie à épifluorescence. Le nombre de cellules fluorescentes et l'intensité moyenne de la fluorescence (m) sont quantifiés par cytométrie de flux.

B. La chromatine des cellules de la lignée stable H1299 BRDT ΔC-GFP est visualisée en microscopie à épifluorescence grâce à un marquage par le Hoechst. Le traitement à la TSA induit un effondrement drastique de la chromatine.

C. Analyse en western blot de la localisation des protéines GFP et BRDT ΔC-GFP après un fractionnement sub-cellulaire de l'échantillon protéique. Les détections de HDAC1 (cytosoluble) et de H3 (histone canonique composant les nucléosome) sont utilisées comme contrôles du fractionnement. Cyto : extrait protéique cytosoluble. Noyau : extrait protéique non cytosoluble.

D. Observation des protéines BRDT ΔC-GFP, p21, H4 acétylée et H3 en Western blot dans les lignées H1299 GFP et H1299 BRDT ΔC-GFP au cours d'une cinétique de traitement par la TSA. Les histones sont visualisées par une coloration au bleu de Coomassie. L'actine est utilisée comme contrôle de charge.

3. Discussion

En accord avec différentes études précédemment publiées (Jones et al., 1997 ; Pivot-Pajot et al., 2003), nos données montrent que l'expression du gène *BRDT*, spécifique du testicule, est réprimée dans le poumon non-tumoral. Mais en tant que gène CT, *BRDT* est fréquemment exprimé de manière aberrante dans les tumeurs somatiques (Jones et al., 1997 ; Scanlan et al., 2002b), y compris dans les tumeurs pulmonaires, et notamment celles des patients de la cohorte « Cancer du poumon ». Dans ce contexte, les travaux préliminaires que j'ai mené sur la protéine BRDT m'ont permis de réunir des outils qui seront utiles pour étudier plus en avant les conséquences de l'expression aberrante de ce facteur testiculaire dans les cancers du poumon. Il s'agit d'une part de l'identification de la lignée H827 de cancer du poumon non à petites cellules comme lignée qui exprime de manière endogène la protéine BRDT, et d'autre part de l'établissement de lignées stables de cellules cancéreuses pulmonaires H1299 exprimant des protéines BRDT fusionnées à la GFP.

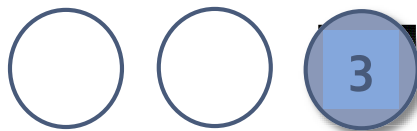
Les cellules des lignées stables H1299 BRDT-GFP, exprimant de manière stable le transgène, exhibent des niveaux d'expression des protéines BRDT proches du niveau d'expression dans les cellules cancéreuses BRDT-positives, ce qui permet de s'affranchir des effets dus à la forte surexpression lors d'une transfection transitoire. Les premiers résultats obtenus dans la lignée H1299 BRDT Δ C-GFP démontrent que la protéine, lorsqu'exprimée hors de son contexte testiculaire, conserve ses propriétés fonctionnelles. Ainsi, la protéine BRDT Δ C-GFP se localise exclusivement dans le noyau des cellules cancéreuses pulmonaires, elle s'associe aux chromosomes mitotiques et elle est capable d'interagir avec l'histone H₄ hyperacétylée. En effet, les protéines de la famille BET se localisent dans le compartiment nucléaire, une propriété attribuée à la présence d'une région conservée agissant comme un signal de localisation nucléaire (Fukazawa and Masumi, 2012). Les protéines BET ont non seulement la capacité d'interagir avec la chromatine grâce à leurs bromodomains qui se lient aux résidus d'histones acétylés (Moriniere et al., 2009), mais elles restent même associées aux chromosomes condensés lors des divisions cellulaires (Dey et al., 2003), une interaction qui serait permise par leur motif B conservé (Garcia-Gutierrez et al., 2012). Nous avons également observé que la protéine BRDT Δ C-GFP apparaît sous la forme de deux entités de poids moléculaires distincts, et nous suggérons que l'une de ces formes corresponde à la protéine phosphorylée. Récemment, une étude a démontré qu'un autre membre de la famille BET, le facteur Brd4 ubiquitaire, est phosphorylé par la caséine kinase II et que cette modification post-traductionnelle conditionne la capacité de Brd4 à interagir avec la chromatine (Wu et al., 2013).

Dans ce modèle de la lignée stable H1299 BRDT Δ C-GFP, nous avons amorcé l'étude des conséquences d'une hyperacétylation des histones. Nous montrons que la TSA induit une accumulation de la protéine BRDT et son association avec la chromatine (puisque la protéine de fusion quitte la fraction soluble et est détectée dans la fraction nucléaire insoluble). Mais avant tout, nous avons montré l'apparition d'une compaction aberrante de la chromatine. Ainsi, dans notre modèle de lignée de cancer pulmonaire qui exprime la protéine BRDT de manière stable et à niveau d'expression modéré, nous retrouvons la capacité de la protéine à induire un effondrement de la structure chromatinienne qui avait précédemment été décrite lors d'une transfection transitoire en cellules COS7 (Pivot-Pajot et al., 2003). Les mécanismes à l'origine de cette compaction n'ont pas été décrits, même si l'on soupçonne la protéine BRDT de pouvoir favoriser une déstabilisation ou même une éviction des histones. Ce phénomène, physiologique dans le contexte de la spermatogenèse où l'ADN s'associe ensuite aux protamines, pourrait avoir des conséquences dramatiques dans une cellule somatique. Dans cette étude préliminaire, nous avons observé une augmentation plus précoce de l'acétylation de H₄ - qui pourrait s'expliquer par un rôle stabilisateur de BRDT sur cette modification - ainsi qu'une perte transitoire du signal de H₃ en western blot, apparemment non corrélée à une perte d'expression de l'histone, et qui pourrait peut-être être due à la présence de modifications post-traductionnelles de H₃. Ces premiers résultats pourraient être complétés par une étude plus globale des niveaux d'expression, modifications post-traductionnelles et localisations des différentes histones canoniques lors de la compaction de la chromatine.

Cette compaction anormale de la chromatine médiée par BRDT suite à l'hyperacétylation des histones est l'observation qui nous a fait émettre notre principale hypothèse de travail : dans les cellules cancéreuses pulmonaires qui expriment le facteur BRDT de manière aberrante, la capacité de BRDT à déstabiliser la structure chromatinienne pourrait être dramatique pour l'homéostasie cellulaire et conduire à la mort de la cellule. En conséquence, les cellules BRDT-positives pourraient s'avérer particulièrement sensibles à une hyperacétylation des histones. Dans ce contexte, j'ai utilisé les deux lignées de cancer du poumon non à petites cellules H827 et H1299, respectivement BRDT-positif et négatif, pour mesurer l'apoptose induite par la TSA. Les données préliminaires obtenues semblent indiquer que les cellules H827 sont plus sensibles à l'apoptose induite par un traitement TSA que les cellules H1299. Néanmoins, la TSA a été décrite comme capable d'induire l'apoptose de cellules cancéreuses par l'activation des voies mitochondriale et des récepteurs de mort (Kim et al., 2006b). Ainsi, la différence de sensibilité des cellules H827 et H1299 à un traitement à la TSA pourrait être indépendante de BRDT et s'expliquer par des différences intrinsèques dans les voies d'apoptose. Pour démontrer l'implication de BRDT dans cette apoptose TSA-induite, nous avons réalisé un silençage de l'expression du gène par interférence ARN. Nos résultats indiquent que la perte d'expression de *BRDT* est associée à une réduction de la mortalité consécutive à un traitement TSA, ce qui suggère une implication directe de BRDT dans la mort TSA-induite. Toutefois, ces résultats n'ont été obtenus que sur une unique expérience et nécessitent d'être reproduits avant de pouvoir conclure définitivement à un rôle de BRDT dans la sensibilité des cellules aux inhibiteurs de HDAC. D'autre part, nous n'avons pas observé de compaction aberrante de la chromatine dans les cellules H827 traitées à la TSA. Si l'implication de BRDT dans la mort TSA-induite se confirme, il faudra alors envisager que l'apoptose observée ne soit pas due à un remodelage drastique de la chromatine hyperacétylée, mais à des effets plus discrets de l'expression ectopique de BRDT.

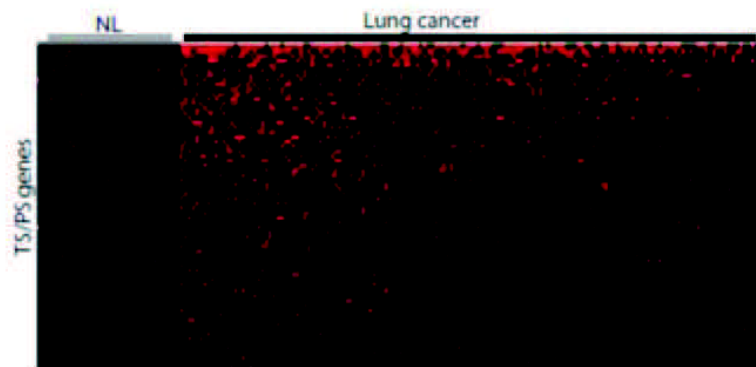
En dehors de son rôle dans la réorganisation physique du génome et l'éviction des histones dans les étapes post-méiotiques de la spermatogenèse, des résultats plus récents démontrent un rôle de BRDT en tant que régulateur transcriptionnel, et ce dès les étapes méiotiques (Gaucher et al., 2012). La protéine BRDT, ciblée en des régions du génome qui portent les marques d'acétylation de H4K5 et K8, est capable de recruter le complexe pTEF β , et induit l'expression d'un répertoire de gènes activés dans les étapes post-méiotiques. On peut alors suggérer que, lorsqu'exprimé de manière aberrante dans les cellules cancéreuses pulmonaires, le facteur BRDT pourrait entraîner l'activation aberrante d'un programme génique spermatogénique et l'expression ectopique de nombreux autres facteurs testiculaires. Cette hypothèse est actuellement explorée au laboratoire grâce à des études en RNA seq sur les lignées H827 et H1299 identifiées dans ce travail.

A l'inverse de ce qui a été observé pour le gène *PRL*, l'expression aberrante de *BRDT* dans les tumeurs pulmonaires n'est pas associée à un pronostic sombre et ne semble pas affecter la survie des patients. Néanmoins, les travaux préliminaires présentés ici soulignent la possibilité d'utiliser l'expression aberrante de BRDT dans une tumeur comme indicateur de réponse à un traitement par les inhibiteurs de HDAC, dont certains, comme le Vorinostat, sont déjà utilisés en clinique pour le traitement des lymphomes notamment. Néanmoins, même si les cellules H827 que nous avons utilisées comme modèles de cellules BRDT-positives sont plus sensibles à un traitement par la TSA que les cellules H1299, il reste le taux d'apoptose atteint est modeste. Il faudrait tester la possibilité de combiner une approche chimiothérapeutique classique avec un inhibiteur de HDAC dans le traitement des tumeurs BRDT-positives. Alors que l'impact de l'expression de BRDT dans les cancers pulmonaires, dont l'étude a été amorcée dans ce travail, reste encore relativement méconnue, l'expression ectopique de BRDT et des autres protéines BET dans les leucémies et lymphomes a davantage été étudiée et participe à la transformation cellulaire (Wu et al., 2013). Dans ces tumeurs non-solides, la protéine Brd4 par exemple agit comme régulateur transcriptionnel, notamment sur le gène *myc*. En conséquence, on assiste au développement d'inhibiteurs spécifiques des bromodomaines (Filippakopoulos et al., 2010 ; Picaud et al., 2013) qui commencent à montrer des résultats encourageants en tant que thérapies anti-cancéreuses dans les néoplasies lymphoïdes (Emadali, 2013 ; Puissant et al., 2013).

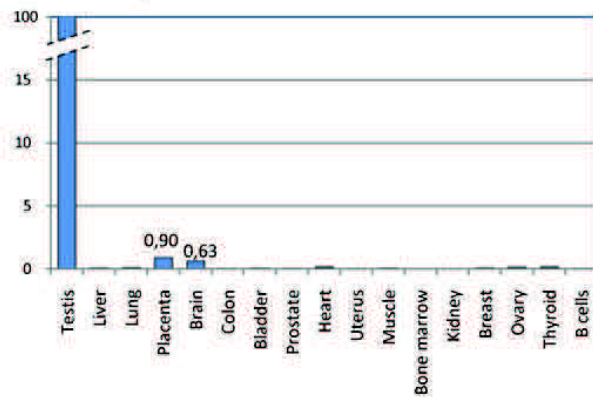


Etude des profils d'expression des gènes *SOX30* et *SPATA22*
exprimés de manière aberrante dans les tumeurs pulmonaires :
De nouveaux facteurs CT ?

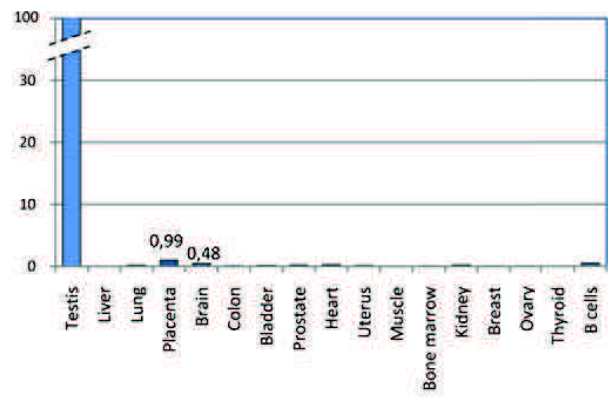
A.



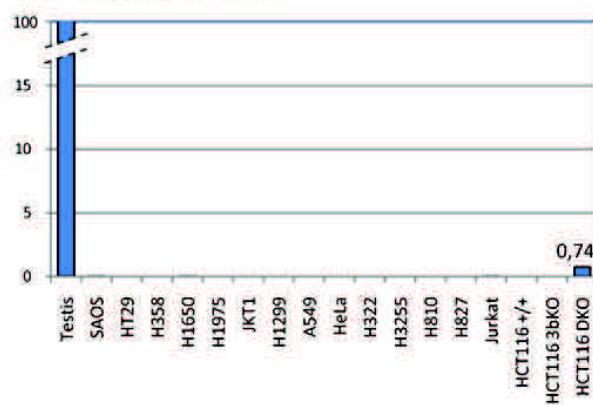
B. *SPATA22* expression ratio



SOX30 expression ratio



C. *SPATA22* expression ratio



SOX30 expression ratio

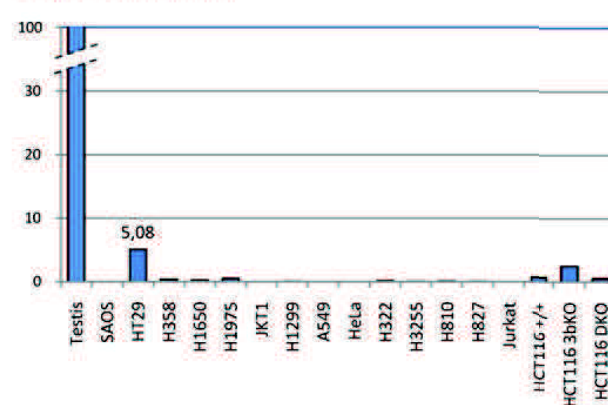


Figure 75 : *SOX30* et *SPATA22*, gènes testiculaires, sont activés dans les tumeurs pulmonaires

A. Tiré de (Rousseaux, 2013). Les gènes *SPATA22* et *SOX30* sont identifiés comme des activations des gènes *SPATA22* et *SOX30* sont fréquemment observées dans les tumeurs de la cohorte de patients « Cancer du poumon ».

B. Profils d'expression des gènes *SPATA22* et *SOX30* obtenus par qPCR dans un panel de tissus humains sains. Les primers utilisés sont spécifiques des 6 variants d'ARN messager de *SPATA22* d'une part et des 2 variants de *SOX30* d'autre part. Normalisation par *Actine* et *U6*. 100 % = Ratio d'expression dans le testicule.

C. Profils d'expression des gènes *SPATA22* et *SOX30* obtenus par qPCR dans un panel de lignées cancéreuses humaines.

1. Contexte de l'étude

Dans le cadre de l'étude systématique des expressions hors-contexte dans la cohorte de patients "Cancer du poumon" récemment publiée ([Rousseaux, 2013](#)), Sophie Rousseaux a mis à jour un groupe de gènes appelés TS/PS (pour Testis-Specific/ Placenta-Specific) dont l'expression est normalement restreinte à la gonade mâle ou au placenta, mais qui sont ré-exprimés dans les tumeurs pulmonaires (**Figure 75A**). Parmi ces gènes, deux candidats fréquemment activés dans les cancers du poumon ont retenu notre attention : le gène *SPATA22* d'une part, de fonction complètement inconnue au moment de l'initiation du projet, et *SOX30* d'autre part, un gène appartenant à la famille SOX qui comprend de nombreux régulateurs transcriptionnels. Les travaux préliminaires que je présente ici visaient à confirmer la spécificité de ces expressions dans le testicule et à préciser les patrons d'expression au sein des cellules de la lignée germinale mâle, mais également à rechercher des lignées cellulaires de cancer de poumon exprimant les gènes *SOX30* et *SPATA22* qui pourraient être utilisées comme modèles pour décrypter les conséquences de l'expression aberrante de ces deux facteurs testiculaires dans les cellules cancéreuses pulmonaires.

2. Résultats

Les gènes *SOX30* et *SPATA22* ont été sélectionnés et renommés TS/PS ([Rousseaux, 2013](#)) du fait de l'analyse de données transcriptomiques qui tendaient à démontrer une expression de ces deux gènes restreinte au testicule. J'ai vérifié ces profils d'expression dans 17 tissus humains par RT-qPCR (**Figure 75B**). Je confirme l'expression prédominante des gènes *SOX30* et *SPATA22* dans le testicule, étant donné que l'expression placentaire n'atteint même pas 1% de l'expression testiculaire. Aucune expression n'est détectée dans le poumon sain.

De fréquentes ré-expressions des gènes TS/PS (incluant *SOX30* et *SPATA22*) ont été observées dans les tumeurs pulmonaires (**Figure 75A**). Néanmoins, lors d'une analyse de l'expression de ces gènes dans des lignées cellulaires cancéreuses, je n'ai pu identifier qu'une lignée qui exprime l'un de ces gènes testiculaires : il s'agit de la lignée de carcinome colorectal HT29 qui exhibe une expression de *SOX30* à 5% de son niveau d'expression testiculaire. Aucune des lignées de cancer du poumon testées ne semble exprimer de manière aberrante l'un ou l'autre de ces gènes.

Les fonctions physiologiques des gènes *SOX30* et *SPATA22* dans les cellules germinales mâles, et plus encore les conséquences de leur expression dans les cellules cancéreuses somatiques, sont inconnues. J'ai donc entrepris une rapide caractérisation de l'expression de ces deux facteurs lors de la spermatogenèse, en établissant les profils d'expression des protéines *SOX30* et *SPATA22* dans les cellules germinales mâles.

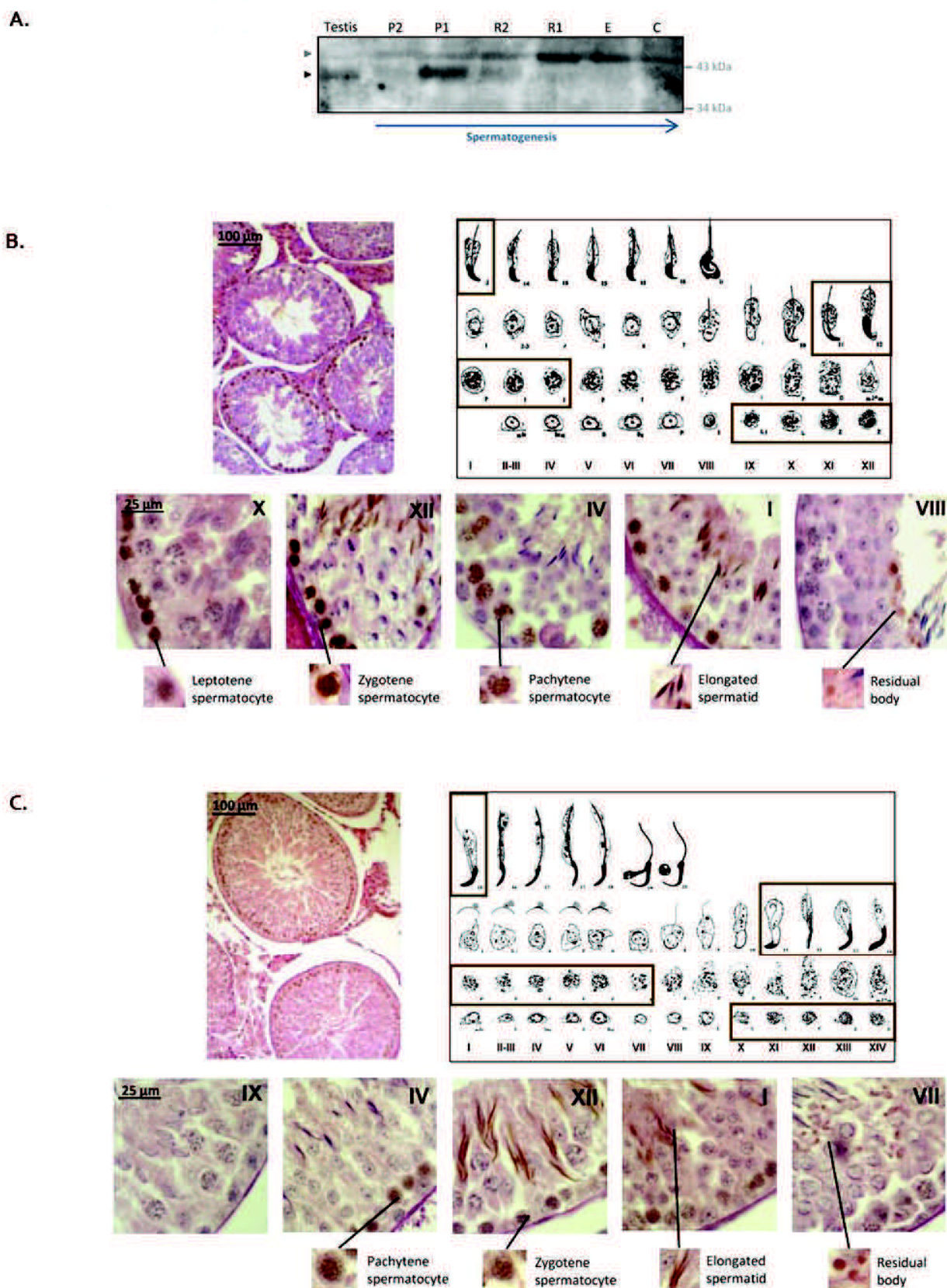


Figure 76: Expression pattern of SPATA22 during male germ cell differentiation in rodent testes

A. Analysis of SPATA22 expression during mouse spermatogenesis. After isolation from mouse testis, germ cells were sorted by sedimentation on BSA gradient. 20 μ g of protein extracts from each fraction were analysed in western blot.

B.C. Transverse histological paraffin-embedded sections of mouse (B) or rat (C) adult testis were immunostained with anti-SPATA22 antibody. DAB staining produces a brown color at the site of antibody localization. Roman numerals within each tubule indicate seminiferous epithelium stages. Upper right hand corner panel shows a summary of the protein localization (brown rectangles), superimposed on the map of spermatogenesis (adapted from (Russell, 1990)).

L: leptotene spermatocyte. Z: zygotene spermatocyte. P: pachytene spermatocyte. D: diplotene spermatocyte. m: meiosis. R: round spermatids. E: elongated spermatids. C: condensed spermatids. Arabic numerals: stages of spermatid differentiation.

2.1. Profil d'expression de SPATA22 dans les cellules spermatogéniques

Lors du lancement de ce projet, rien n'était décrit sur la protéine SPATA22 (spermatogenesis associated 22) dont la fonction demeurerait complètement inconnue. A l'exception de la description de variants d'épissage alternatifs dans le testicule (Huang et al., 2005), ni le patron d'expression du gène ni celui de la protéine n'étaient connus.

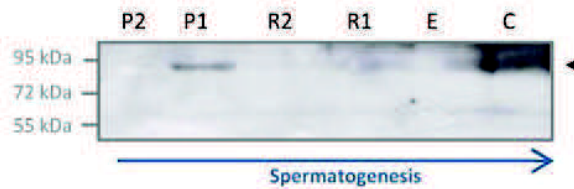
Nous avons étudié le profil d'expression de SPATA22 durant la spermatogenèse murine (**Figure 76A**) en western blot sur les différentes populations cellulaires germinales. Un anticorps anti-SPATA22 reconnaît une bande à la taille attendue (40 kDa, flèche noire) principalement dans les spermatocytes, mais marque également une bande de poids moléculaire plus élevé (45 kDa, flèche grise) dans les spermatides rondes et élonguées correspondant aux stades plus tardifs de la spermatogenèse. Etant donné qu'aucune autre isoforme protéique de SPATA22 n'a été décrite, nous proposons que cette bande de taille apparente plus élevée puisse correspondre à une protéine SPATA22 modifiée, par exemple par une phosphorylation sur le résidu Y275 décrit comme potentiellement phosphorylable.

Dans un second temps, nous avons souhaité préciser ce profil d'expression spermatogénique en localisant la protéine SPATA22 par immunohistochimie sur des coupes de testicules de rongeurs (**Figures 76B et C**). Chez la souris (**Figure 76B**), une forte immunoréactivité (marquage brun) révèle que SPATA22 est fortement exprimé durant la phase méiotique, dans les noyaux des spermatocytes leptotènes (stades VIII à X), zygotènes (XI et XII) et pachytènes précoces (stades I à IV). Aucun signal n'est détecté dans les spermatogonies (stades II à VII), dans les spermatocytes pachytènes tardifs et diplotènes (stades V à XII) ou dans les spermatides rondes ou en élongation (stades I à X). De manière intéressante, SPATA22 est détectée plus tardivement au cours de la phase post-méiotique : les noyaux des spermatides élonguées/condensées sont marqués entre les stades 11 à 13 (XI à I). Alors que les stades finaux de la différenciation des spermatides n'exhibent plus de signal, on observe le marquage des corps résiduels (au stade VIII). Ce profil d'expression chez la souris est hautement semblable à celui qui peut être observé chez le rat (**Figure 76C**). De la même manière, SPATA22 est détectée au cours des phases méiotiques et post-méiotiques, dans les noyaux des spermatocytes leptotènes (X à XII), zygotènes (XII et XIII), pachytènes (XIV to VII), puis dans les noyaux des spermatides des stades 11 à 15 (XI à I) ainsi que dans les corps résiduels.

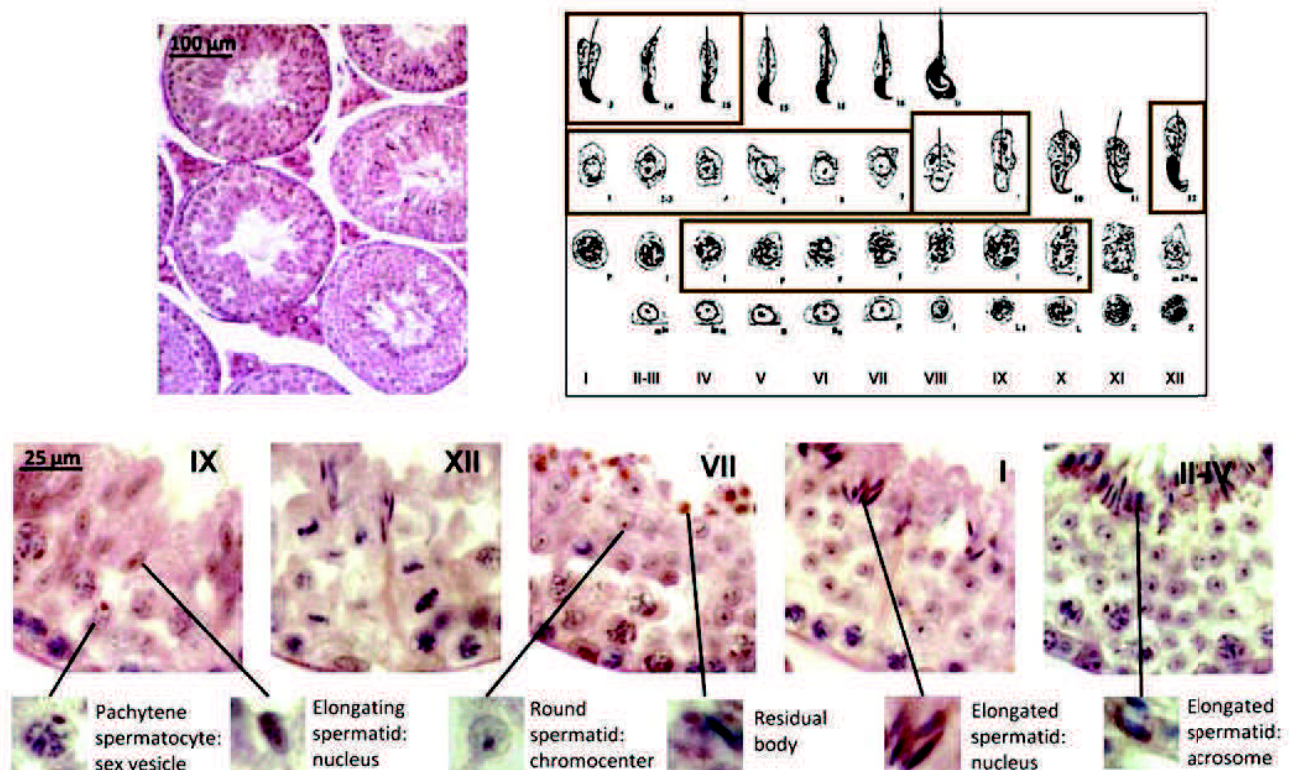
2.2. Profil d'expression de SOX30 dans les cellules spermatogéniques

La même approche a été suivie pour étudier le profil d'expression de SOX 30 (sex determining region Y-box 30) au cours de la spermatogenèse chez la souris. En western blot, la protéine SOX30 est détectée à la taille attendue de 80 kDa dans la fraction P1 dans spermatocytes pachytènes, très légèrement dans les spermatides rondes et en élongation, et de manière plus intense dans les spermatides condensées (**Figure 77A**). En immunohistochimie sur coupe de testicule murin, aucun marquage n'est observé dans les spermatogonies et les spermatocytes précoces (**Figure 77B**). Dans les spermatocytes pachytènes tardifs (stades IV à X), un marquage SOX30 intense est localisé dans la vésicule sexuelle. Dans les spermatides rondes (stades I à VII), SOX30 se localise dans une zone ponctuée proche du chromocentre. Aux stades VIII et IX, le marquage s'étend à l'ensemble du noyau des spermatides en cours d'élongation. Ce marquage est intense dans les spermatides élonguées des stades 12 à 14 (XII à I). Dans les étapes tardives de la différenciation des spermatides, SOX30 n'est plus visible dans le noyau mais on détecte un marquage de l'acrosome (stade IV) et des corps résiduels (stade VII). Pour préciser la localisation dans la vésicule sexuelle et le chromocentre, on a réalisé des immunofluorescences sur empreintes de testicule de souris (**Figure 77C**). Un signal est observé dans les acrosomes (colonne 1) et les noyaux (colonne 2) des spermatides condensées. Seul un marquage diffus est observable dans les spermatocytes et les spermatides rondes (colonne 3).

A.



B.



C.

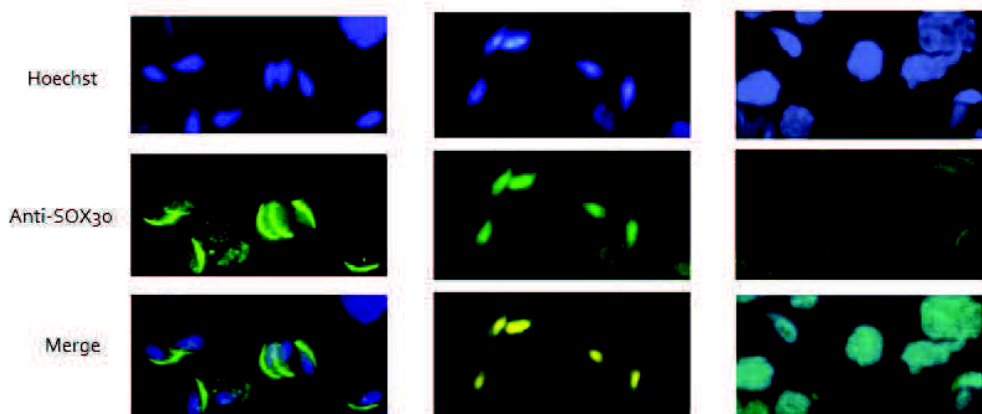


Figure 77: Expression pattern of SOX30 during male germ cell differentiation in mouse testes

A. Analysis of SOX30 expression during mouse spermatogenesis. After isolation from mouse testis, germ cells were sorted by sedimentation on BSA gradient. 20 µg of protein extracts from each fraction were analysed in western blot.

P: spermatocytes. R: round spermatids. E: elongated spermatids. C: condensed spermatids.

B. Transverse histological paraffin-embedded sections of mouse adult testes were immunostained with anti-SOX30 antibody. DAB staining produces a brownish color at the site of antibody localization. Roman numerals within each tubule indicate seminiferous epithelium stages. Upper right panel shows a summary of the protein localization (brown rectangles), superimposed on the map of spermatogenesis (adapted from (Russell, 1990)).

L: leptotene spermatocyte. Z: zygotene spermatocyte. P: pachytene spermatocyte. D: diplotene spermatocyte. m: meiosis. Arabic numerals: stages of spermatid differentiation.

C. Imprinted testicular murine cells were analyzed by immunofluorescence with anti-SOX30 antibody. DNA is stained with Hoeschst.

3. Discussion

Les travaux préliminaires réalisés sur les gènes *SOX30* et *SPATA22* nous ont permis de confirmer l'expression ultra-majoritaire de ces deux gènes dans les cellules de la lignée germinale mâle. Nous avons noté une faible expression de ces gènes dans le placenta. Cette expression placentaire n'avait pas été détectée en Northern blot dans l'étude ayant caractérisée *SOX30* (Osaki et al., 1999) mais est en accord avec les données EST disponibles pour *SPATA22* et *SOX30* où le placenta apparaît comme un site secondaire d'expression des gènes. Cette observation rapproche *SOX30* et *SPATA22* de nombreux facteurs CT pour lesquels une expression placentaire s'ajoute à l'expression testiculaire. Ce profil d'expression des gènes *SOX30* et *SPATA22* restreint au testicule et au placenta et leur fréquente expression dans les tumeurs pulmonaires nous permettent de proposer *SOX30* et *SPATA22* comme de nouveaux facteurs CT. On peut noter que le gène *SPATA19*, qui pourrait être impliqué dans la spermiogenèse et dont l'expression a été décrite dans des cancers de la peau et de la prostate (Ghafouri-Fard et al., 2010), fait partie de la liste actuelle des facteurs CT sous l'identifiant CT132.

Quelques données sont disponibles sur le facteur *SOX30* dans la littérature. Il a été identifié comme une nouvelle protéine à domaine high mobility group (HMG)-box et a été inclus dans la famille des gènes SOX (SRY-related HMG box) (Osaki et al., 1999). Les gènes SOX, très conservés au cours de l'évolution, sont impliqués dans de nombreux processus développementaux, et de nombreuses protéines SOX régulent l'expression des gènes en agissant comme des activateurs ou des répresseurs transcriptionnels (Kiefer, 2007). *SOX30* lui-même semble capable d'activer la transcription d'un promoteur synthétique (Osaki et al., 1999). Ce potentiel rôle dans le contrôle des expressions géniques nous permet d'imaginer que *SOX30*, exprimé de manière aberrante dans les cellules cancéreuses pulmonaires, puisse activer un programme d'expression génique. Mais jusqu'à aujourd'hui, aucune étude ne s'est penchée sur les conséquences de l'expression de *SOX30* dans les cellules cancéreuses.

Alors que rien n'était connu sur la fonction assurée par *SPATA22* dans le testicule lors de l'obtention de ces résultats, une étude récente (La Salle et al., 2012) a observé qu'une mutation de *SPATA22* cause une infertilité, par blocage de la spermatogenèse dès le stade méiotique. Cette découverte est en accord avec notre observation de la protéine *SPATA22* dans les cellules méiotiques et permet de suggérer un rôle important de *SPATA22* dans le bon déroulement de la spermatogenèse. Quant aux conséquences de l'expression de *SPATA22* dans des cellules cancéreuses somatiques, elles restent encore à explorer.

CONCLUSION

Conclusion générale

Mon travail de thèse s'est inscrit dans un projet innovant initié par le Docteur Sophie Rousseaux, basé sur une approche translationnelle de recherche systématique des expressions hors-contexte de gènes tissu-spécifiques dans les cancers pulmonaires. Notre travail est basé sur l'hypothèse que ces activations ectopiques, bien moins connues et étudiées que les répressions aberrantes de gènes suppresseurs de tumeurs, pourraient pourtant constituer une source importante et innovante de facteurs impliqués dans la tumorigenèse et de potentielles cibles thérapeutiques. Mes travaux ont visé à caractériser les causes et les conséquences des activations ectopiques de quelques candidats, les gènes *BRDT*, *SOX30*, *SPATA22* et plus particulièrement le gène *PRL* codant la prolactine. Parmi l'ensemble des activations aberrantes mises à jour dans les cancers pulmonaires, ces gènes candidats ont été retenus pour une étude plus poussée pour des raisons distinctes. Les gènes *SPATA22* et *SOX30*, dont les fonctions sont méconnues, exhibent un profil d'expression restreint au testicule, mais sont fréquemment activés dans les cancers pulmonaires, ce qui peut nous conduire à les proposer comme nouveaux gènes CT. Notre intérêt pour l'expression hors-contexte du facteur *BRDT* a été guidé par les connaissances accumulées sur son rôle d'effecteur épigénétique puissant dans la lignée germinale mâle et s'est appuyé sur l'hypothèse que la fonction de *BRDT* pourrait être « retournée » contre les cellules cancéreuses grâce à un traitement par HDACi. Enfin la réexpression aberrante du gène hypophysaire *PRL* dans les cellules cancéreuses pulmonaires a retenu notre attention du fait d'une forte association avec un pronostic sombre pour le patient. Les résultats obtenus durant mon travail de thèse, depuis des analyses de données cliniques jusqu'à des approches biochimiques dans des modèles cellulaires, illustrent l'intérêt de l'étude de ces activations aberrantes. En particulier, l'expression ectopique du gène *PRL* se fait dans un cadre totalement inédit de production d'un messager tronqué, où la protéine non-sécrétée pourrait agir indépendamment de son récepteur pour promouvoir la résistance et l'agressivité des cellules cancéreuses pulmonaires. L'approche que nous avons suivie souligne l'importance de la régulation fine des expressions géniques, et met en lumière les conséquences dramatiques que peuvent avoir ces expressions ectopiques. Mes travaux illustrent également le fait que ces expressions ectopiques constituent un réservoir de potentiels facteurs oncogènes, biomarqueurs de tumeurs ou de pronostic, dont l'étude plus poussée pourrait permettre de proposer de nouvelles approches thérapeutiques ciblées, dans le cancer du poumon mais aussi dans l'ensemble des pathologies néoplasiques caractérisées par un profond bouleversement des expressions géniques.

MATERIELS ET METHODES

Nom	Provenance	En suspension	Adhérentes	Milieu complet utilisé	Dilution de passage
H1299	Carcinome pulmonaire (CBNPC)		X	RPMI complet	1/8
A549	Carcinome pulmonaire (CBNPC)		X	RPMI complet	1/10
H810	Carcinome pulmonaire (CBNPC)		X	Milieu HITES = RPMI complet + 0,005 mg/mL insuline + 0,01 mg/mL transferrine + 30 nM NaSel + 10 nM hydrocortisone + 10 nM estradiol + 10 mM HEPES	1/3
H827	Adénocarcinome pulmonaire (CBNPC)		X	RPMI complet	1/5
H1650	Adénocarcinome pulmonaire (CBNPC)		X	RPMI complet	1/5
H146	Carcinome pulmonaire (CBPC)	X		RPMI complet + 10 mM HEPES + 1 mM NaPyr + 4,5 g/L glucose + 1,5 g/L NaBic	1/2
H524	Carcinome pulmonaire (CBPC)	X		RPMI complet	1/3
H69	Carcinome pulmonaire (CBPC)	X		RPMI complet + 10 mM HEPES + 1 mM NaPyr + 2,5 g/L glucose	1/3
H526	Carcinome pulmonaire (CBPC)	X		RPMI complet	1/3
HCC2429	Cancer de la ligne médiane à translocation BRD4-NUT		X	RPMI complet	1/5
HeLa	Adénocarcinome cervical		X	DMEM 1,1 g/L glucose complet	1/10
HCT116	Carcinome colorectal		X	Mc Coy complet	1/8
U2OS	Ostéosarcome		X	Mc Coy complet	1/5
Jurkat	Lymphome à cellules T	X		RPMI complet	1/8
COS	Rein de singe vervet		X	DMEM 1,1 g/L glucose complet	1/8

Figure 78 : Lignées cellulaires utilisées : provenance, caractéristiques et conditions de culture

Toutes les lignées cellulaires utilisées sont humaines, à l'exception des cellules COS.

Le terme « milieu complet » désigne un milieu complémenté de 10% SVF, 2% L-glutamine et antibiotiques.

1. MATERIELS

1.1. Matériels biologiques

Les lignées cellulaires utilisées ont été obtenues auprès de l'American Type Culture Collection (ATCC), à l'exception de la lignée HCT116 DKO, double KO pour les ADN méthyltransférases DNMT1 et DNMT3b, précédemment caractérisée ([Rhee et al., 2002](#)) qui a été fournie par le Dr Volgestein (Johns Hopkins, Baltimore, MD).

Les souris OF1 (Oncins France 1) ont été commandées auprès de la société Charles River (Elbeuf, France).

Les ARN de tissus humains ont été obtenus auprès de la société Clontech (Human Total RNA, Master Panel II). Les ARN de placenta ont été obtenus auprès du Dr Nadia Alfaidy-Benharouga (CEA, Grenoble).

Les plasmides BRDT (BRDT Δ C-GFP et HA-BRDT FL-GFP) ont été obtenus auprès de Jonathan Gaucher ([Gaucher et al., 2012](#)). Les plasmides NUT (GST-NUT f1c et GFP-BRD4-NUT) ont été obtenus auprès de Nicolas Reynoird ([Reynoird et al., 2010b](#)).

1.2. La cohorte de patients "Cancer du poumon"

La cohorte "Cancer du Poumon" compte 293 patients ayant été suivis pour une tumeur pulmonaire au Centre Hospitalier Universitaire de Grenoble et ayant subi une chirurgie. Pendant plus de 10 ans, les Professeurs Elisabeth Brambilla et Christian Brambilla ont recueilli des données cliniques (âge du patient, sexe du patient, statut tabagique, date du diagnostic, classification histologique de la tumeur, classification TNM de la tumeur, traitements subis, date du décès...) et des échantillons biologiques (biopsies, coupes de tumeurs, ARN). Pour chaque patient, une étude transcriptomique « genome-wide » a été réalisée sur un échantillon de tumeur prélevé avant mise en place des traitements, et dans certains cas sur un échantillon de tissu pulmonaire non-cancéreux adjacent. Les données transcriptomiques ont été obtenues avec la technologie Affymetrix ("Affymetrix.GeneChip.HG-U133_Plus_2"). Ces données sont disponibles sur le site GEO sous le référencement GSE30219 (**Voir paragraphe 2.4.1** pour les méthodes d'analyse transcriptomique).

Utilisation ultérieure	Contenant	Surface (cm ²)	Milieu (mL)	Ensemencement pour récolte à J+3	Ensemencement pour récolte à J+2
Extraction de protéines	Boîte Ø100	57	10	500 000 cellules	1 000 000 cellules
Extraction d'ARN	Boîte Ø60	20,4	6	200 000 cellules	400 000 cellules
Analyse en FACS	Plaquette 6 puits	9,6	4	100 000 cellules	200 000 cellules
Analyse sur lames	Labteks 2 puits	4	2	40 000 cellules	80 000 cellules

Figure 79 : Conditions d'ensemencement des lignées cellulaires cancéreuses selon l'utilisation ultérieure

2. METHODES

2.1. Méthodes sur cellules et tissus

2.1.1. Purification et fractionnement des cellules germinales mâles

Une souris OF1 mâle est sacrifiée par dislocation cervicale et on prélève les testicules. On réalise un fractionnement des différentes cellules germinales (spermatocytes, spermatides rondes et spermatides élonguées/condensées) par sédimentation (Gaucher et al., 2012). Brièvement, le testicule est prélevé, lavé dans le PBS et débarassé de son albuginée. Les tubules séminifères sont libérés par une incubation de 10 minutes à 35°C dans une solution de collagénase (Sigma) diluée à 1 mg/mL dans du tampon phosphate salin (PBS). Les cellules germinales sont extraites des tubules par pipettage dans le milieu de sédimentation DMEM (Dulbecco's modified Eagle's medium) (Gibco) additionné de 0,5% de BSA (Euromedex) et de 2 µg/mL de DNase (Sigma). Une fois la suspension homogène, elle est filtrée deux fois à l'aide de filtres à pores de 100 µm de diamètre (Becton Dickinson). Les cellules, resuspendues dans 18 mL du milieu de sédimentation, sont déposées sur un gradient de BSA de 360 mL contenu dans une chambre de sédimentation. Le gradient est obtenu en remplissant lentement la chambre avec deux solutions de 180 mL à 2 et 4% de BSA diluée dans du DMEM, à l'aide d'un système de robinets permettant le mélange des deux solutions avant leur entrée dans la chambre. Après sédimentation des cellules pendant 70 minutes à 4°C, des fractions de 10 mL sont récoltées sur glace puis centrifugées (1000 rpm pendant 10 minutes à 4°C). Les cellules de chaque fraction sont identifiées sous microscope optique et regroupées pour obtenir des fractions cellulaires de chaque stade enrichies à 80% (spermatocytes pachytènes, spermatides rondes, spermatides allongées et spermatides condensées). Ces différentes populations cellulaires sont ensuite utilisées pour extraction de protéines (voir paragraphe 2.3.1.1).

2.1.2. Culture de lignées cellulaires cancéreuses

2.1.2.1. Lignées cellulaires utilisées et conditions de culture

Les cellules cancéreuses utilisées, leurs provenances, leurs caractéristiques, les milieux de culture complétés utilisés et les dilutions de passage sont listés dans le tableau de la page précédente (**Figure 78**). Les cellules sont cultivées dans un incubateur à 37°C, 5% de CO₂, 95 % air atmosphérique et 90 % d'hygrométrie. Les milieux de culture RPMI (Roswell Park Memorial Institute medium) 1640 (Gibco), DMEM (Dulbecco's modified Eagle's medium) (Gibco) et Mc Coy (Gibco) sont complétés de 10% sérum de veau fœtal (SVF) (PAA), de 2% L-glutamine (200 mM, Gibco) et d'antibiotiques (Pénicilline, streptomycine, 100X, Gibco). Ils peuvent être supplémentés de sodium pyruvate (Gibco), de sodium bicarbonate (Gibco), d'HEPES (Gibco) ou de glucose (PAA) aux concentrations indiquées. Ils sont conservés 1 mois à 4°C.

Les passages des cellules adhérentes sont réalisés les lundis et vendredi, lorsque la confluence atteint 90%, sous un poste de sécurité microbiologique (PSM) allumé à l'avance pour stabiliser le flux d'air : les cellules sont lavées au PBS 1X (10X, PAA), décollées 5 à 10 minutes dans 1 mL d'un mélange Trypsine-EDTA (10 mM, PAA), reprises et dissociées dans du milieu de culture et centrifugées 5 minutes à 1000 rpm pour éliminer l'excès de trypsine, puis diluées en boîte de culture de diamètre 100 mm dans du milieu complet préchauffé à 37°C au bain-marie. Les cellules en suspension sont cultivées en flasques de surface 75 cm² et diluées dans du milieu de culture préchauffé les vendredis seulement. On tente de limiter au maximum la dissociation des cellules de CBPC poussant en amas en suspension, étant donné que la rupture des contacts cellulaires induit une forte mortalité.

A.

Nom	Détail de localisation	Séquence brin forward	Séquence brin reverse
Scramble	Contrôle négatif	AGGUAGUGUAAUCGCCUUGTT	AACAAGGCGAUUACACUACCU
siPRL n°1	Position 1173 (Exon 5)	GCCGAAUCAUCCACAACAA	UUGUUGUGGAUGAUUCGGC
siPRL n°2	Position 711 (Exons 2/3)	GUUCAGCGAAUUCGAUAAA	UUUAUCGAAUUCGUGAAC
siPRL-R n°1	Position 893 (Variants 1, 2, 3, 4, 5 et 7)	GUUUCUCGGAUGAACUUUATT	UAAAGUUCAUCCGAGAACTT
siPRL-R n°2	Position 1053 (dans la partie commune aux 7 variants)	GCUCCUGUAUGAAAUUCGAUU	UCGAAUUUCAUACAGGAGCGU
siBRDT	Position 1366	CGGCAGAUGUUAGAUUAAUUU	AUUAAUUCUACAUCUGCCGCU

B.

Contenant	Surface (cm²)	Milieu (mL)	OptiMEM (µL)	Lipofectamine RNAiMax (µL)	Duplex siRNA à 10 mM (µL)
Boîte Ø60	20,4	6	1 mL	10 µL	3 µL
Plaque 6 puits	9,6	4	500 µL	5 µL	1,5 µL
Labteks 2 puits	4	2	250 µL	2,5 µL	0,75 µL

C.

Contenant	Surface (cm²)	Milieu (mL)	OptiMEM (µL)	Lipofectamine 2000 (µL)	ADN plasmidique (µg)
Boîte Ø100	57	10	2 x 1000	10	5
Boîte Ø60	20,4	6	2 x 400	4	2
Plaque 6 puits	9,6	4	2 x 200	2	1
Labteks 2 puits	4	2	2 x 100	1	0,5

Figure 8o : Protocoles de transfection des cellules cancéreuses

A. Liste des siRNA utilisés et séquences correspondantes

B. Conditions de transfection d'un siRNA avec le Lipofectamine RNAiMax (Invitrogen) selon le contenant.

C. Conditions de transfection d'un ADN plasmidique avec la Lipofectamine 2000 (Invitrogen) selon le contenant.

La confluence des cellules, ainsi que l'absence de contamination, sont vérifiées quotidiennement. Le nombre de passage des cellules ne dépasse pas 30. Une fois ce nombre de passages atteint, de nouvelles cellules - à nombre de passage inférieur, conservées dans l'azote liquide à une concentration de 2 à 5 millions de cellules par mL et en présence de diméthylsulfoxyde DMSO (Sigma) - sont décongelées et placées en culture.

2.1.2.2. Ensemencements de cellules cancéreuses

Les cellules cancéreuses sont encemencées différemment selon l'utilisation qui va suivre (analyse en fluorescence, extraction d'ARN ou de protéines...). Ces conditions d'ensemencement particulières (contenant, nombre de cellules) sont décrites dans le tableau de la page précédente (**Figure 79**). Le comptage des cellules est réalisé sous microscope photonique, à l'aide d'une cellule de Malassez. La mortalité est estimée à l'aide d'une coloration par le bleu Trypan (Gibco). Il faut noter que l'ensemencement se fait toujours en milieu complet, sauf dans le cas où les cellules seront transfectées avec des siRNA : le milieu d'ensemencement est alors privé d'antibiotique.

2.1.3. Transfections de cellules cancéreuses en culture

2.1.3.1. Transfection de siRNA

Les siARN sont désignés en utilisant les outils en ligne BLOCK-iT™ RNAi Designer (Invitrogen, <http://rnaidesigner.invitrogen.com/>) ou DSIR (Designer of Small Interfering RNA) (<http://biodev.extra.cea.fr/DSIR/>), à l'exception du siPRL-R n°1 dont la séquence est tirée de la littérature (Pana, 2007, Inhibition of cell proliferation by siRNA targeting hPRLR in breast cancer MCF-7 cell line). Les duplex de siARN sont synthétisés par la société Eurogentec. Les différents siRNA utilisés au cours de cette thèse sont listés dans le tableau ci-contre (**Figure 80A**). La transfection des siRNA se fait de manière concomitante avec l'ensemencement (transfection «forward»). Elle est médiée par le système de transfection par agent lipidique cationique Lipofectamine RNAiMax (Life Technologies), selon les recommandations du fournisseur et en adaptant les quantités à la taille du contenant (**Figure 80B**). Pour un ensemencement en boîte Ø60, le duplex de siRNA (6 pmol, soit 5 µL à 10 mM) est dilué dans 1 mL d'optiMEM et 10 µL de Lipofectamine RNAi. Après incubation de 15 minutes, 400 000 cellules reprises dans du milieu sans antibiotique sont ajoutées sur ce mix. La transfection contrôle est réalisée avec un duplex de siRNA « scramble » (**Figure 80A**). Les cellules sont récoltées après 48h de silençage.

2.1.3.2. Transfection d'ADN plasmidique

La transfection par ADN plasmidique est réalisée le lendemain de l'ensemencement (J+1), grâce au système de transfection par agent lipidique cationique Lipofectamine 2000 (Life Technologies), selon les recommandations du fournisseur, avec un rapport Lipofectamine 2000 : ADN de 2 et en adaptant les quantités à la taille du contenant (**Figure 80C**). Pour une boîte Ø100, 5 µg de vecteur sont dilués dans 1 mL d'OptiMEM d'une part et 10 µL de Lipofectamine sont dilués dans 500 µL d'OptiMEM d'autre part. Après 5 minutes d'incubation, ces deux préparations sont mélangées. Le tout est ajouté sur les cellules en culture après 20 minutes. La transfection contrôle est réalisée avec la même quantité de vecteur vide. Les cellules sont récoltées après 24h de transfection.

Pour l'établissement d'une lignée stable, une pression de sélection de 600 µg/mL de géiniticine (G418, Gibco) est appliquée deux jours après la transfection et est maintenue environ 10 jours, jusqu'à la mort de l'ensemble des cellules d'une boîte contrôle non transfectée. Les cellules sélectionnées sont ensuite maintenues en culture en présence de 300 µg/mL de géiniticine. Les cellules fluorescentes sont ensuite triées en cytométrie de flux (**voir paragraphe 2.1.5.2.3**).

Drogue	Fonction	MM (g/mol)	Stock	Utilisation	
				Concentration	Durée
Actinomycine D (Sigma)	inhibiteur de la RNA polymérase II	1255,42	100 µg/mL dans H ₂ O = 79,65 µM (4°C)	250 nM	16h
C646 = CM 01-28 (Phil Cole, David Meyers)	inhibiteur de p300	445,42	20 mM (-20°C)	30 µM	6h
Géniticine = G418 (Gibco)	Antibiotique de sélection		50 mg/mL (4°C)	600 µg/mL puis 300 µg/mL	Environ une semaine Pour toujours
Sodium butyrate	inhibiteur des HDAC 1 à 5	110,10	2 M pH=7,4 (4°C)	10 mM	6h
Trichostatine A (TSA) (Sigma)	inhibiteur des HDAC 1 à 6	302,40	DMSO : 1 mg/mL H ₂ O : 25 µg/mL = 82,7 µM (-20°C)	100 ng/mL	6h - 24h

Figure 81 : Traitements des cellules cancéreuses en culture

Les fonctions, les caractéristiques, les doses et les durées de traitement des différentes drogues utilisées sont indiquées dans ce tableau.

MM : Masse moléculaire.

2.1.4. Traitements des cellules cancéreuses en culture

Le lendemain de l'ensemencement (J+1), ou le jour suivant la transfection (J+2), les cellules peuvent être traitées par différentes drogues. Vous trouverez les différentes drogues utilisées, leur mode d'action, leurs concentrations et la durée du traitement listées dans le tableau ci-contre (**Figure 81**). Le point contrôle est traité par une quantité égale du milieu de dilution de la drogue (H₂O ou DMSO). Les cellules sont récoltées à l'issue de la durée de traitement.

2.1.5. Immunomarquages

2.1.5.1. Marquage immunofluorescent sur cellules cancéreuses

Les cellules, ensemencées en labteks, sont fixées 5 minutes avec 500 µL de paraformaldéhyde (PFA) 4%/PBS après aspiration du milieu de culture, puis perméabilisées 1 minute avec le même volume de Triton 0,1%/PFA 4%/PBS. Après 3 lavages de 5 minutes au PBS, les sites non-spécifiques sont saturés avec 1 mL de PBS/Lait 5% pendant 30 minutes. L'incubation avec l'anticorps primaire, à la dilution appropriée (**Figure 83**) est réalisée sur la nuit à 4°C dans 100 µL de PBS/Lait 1%. Après 3 lavages de 5 minutes au PBS, les cellules sont mises en présence pendant 30 minutes à 37°C de l'anticorps secondaire couplé à un fluorochrome (anti-souris, anti-rat ou anti-lapin ; Alexa Fluor, Invitrogen) dilué au 1/500 dans 100 µL de lait 1%/PBS. L'excès d'anticorps est supprimé par 3 lavages de 5 minutes au PBS. Les noyaux sont marqués pendant 5 minutes par un intercalant de l'ADN fluorescent, le Hoechst 33258 (1mg/mL) (Sigma-Aldrich), dilué au 1/500 dans 100 µL de PBS. Les cellules ainsi marquées sont recouvertes d'une lamelle en utilisant une goutte d'une solution de montage « anti-fading » (Dako Mounting Medium). En parallèle, le contrôle négatif est réalisé avec incubation de l'anticorps secondaire seul, permettant d'évaluer le bruit de fond dû à l'accrochage non spécifique de cet anticorps. Dans le cas d'une co-immunodétection, les deux anticorps primaires dans un premier temps, puis les deux anticorps secondaires (si d'espèces différentes) dans un second temps, sont mélangés et incubés conjointement avec les cellules. L'observation est réalisée sur un microscope à épifluorescence avec apotome (Imager.Z1 ApoTome AxioCam, Zeiss) grâce au logiciel AxioVision.

2.1.5.2. Cytométrie de flux

2.1.5.2.1. Analyse de la mortalité cellulaire en cytométrie de flux

Le marquage des cellules apoptotiques est réalisé grâce au kit « PE Active Caspase 3 Apoptosis Kit » (BD Pharmingen) selon les recommandations du fournisseur. Les cellules ainsi que le milieu de culture sont récoltés dans des tubes de cytométrie à fond rond par simple pipettage pour les cellules en suspension, par décollement à la trypsine pour les cellules adhérentes. Les cellules sont centrifugées (3000 rpm 3 minutes), dissociées, fixées et perméabilisées simultanément dans 500 µL de tampon « Cytofix/Cytoperm » pendant 20 minutes à 4°C. Après 3 lavages avec 500 µL de tampon « Wash » 1X (et dissociation mécanique des amas par pipettage pour les cellules de CBPC), les cellules sont incubées avec 15 µL d'anticorps anti-caspase 3 activée couplé à la phycoérythrine (PE) dilués dans 100 µL de tampon « Wash » pendant 30 minutes. Les cellules sont à nouveau lavées trois fois, puis reprises dans 1 mL de PBS additionné de Hoechst dilué au 1/500. En parallèle, un tube contrôle est réalisé sans incubation avec l'anticorps. Si la fluorescence rouge est mesurée conjointement avec une fluorescence verte (présence d'une protéine fusionnée à la GFP par exemple), on prépare un tube non marqué, et deux tubes marqués avec un seul fluorochrome : vert seul (pas d'incubation avec l'anticorps anti-caspase 3-activée couplé à la PE) et rouge seul (cellules-mères non transfectées, n'exprimant pas la GFP).

La lecture de la fluorescence est réalisée sur un cytomètre de flux LSRII (BD Biosciences) possédant les lasers permettant l'excitation du Hoechst d'une part, de la GFP et de la PE d'autre part. L'analyse est réalisée grâce au logiciel Diva6 (BD) (**Figure 82**). La population est sélectionnée grâce à une fenêtre dans le graphe présentant le SSC en ordonnées et le FSC en abscisses. Les éventuels doublets sont éliminés en ne

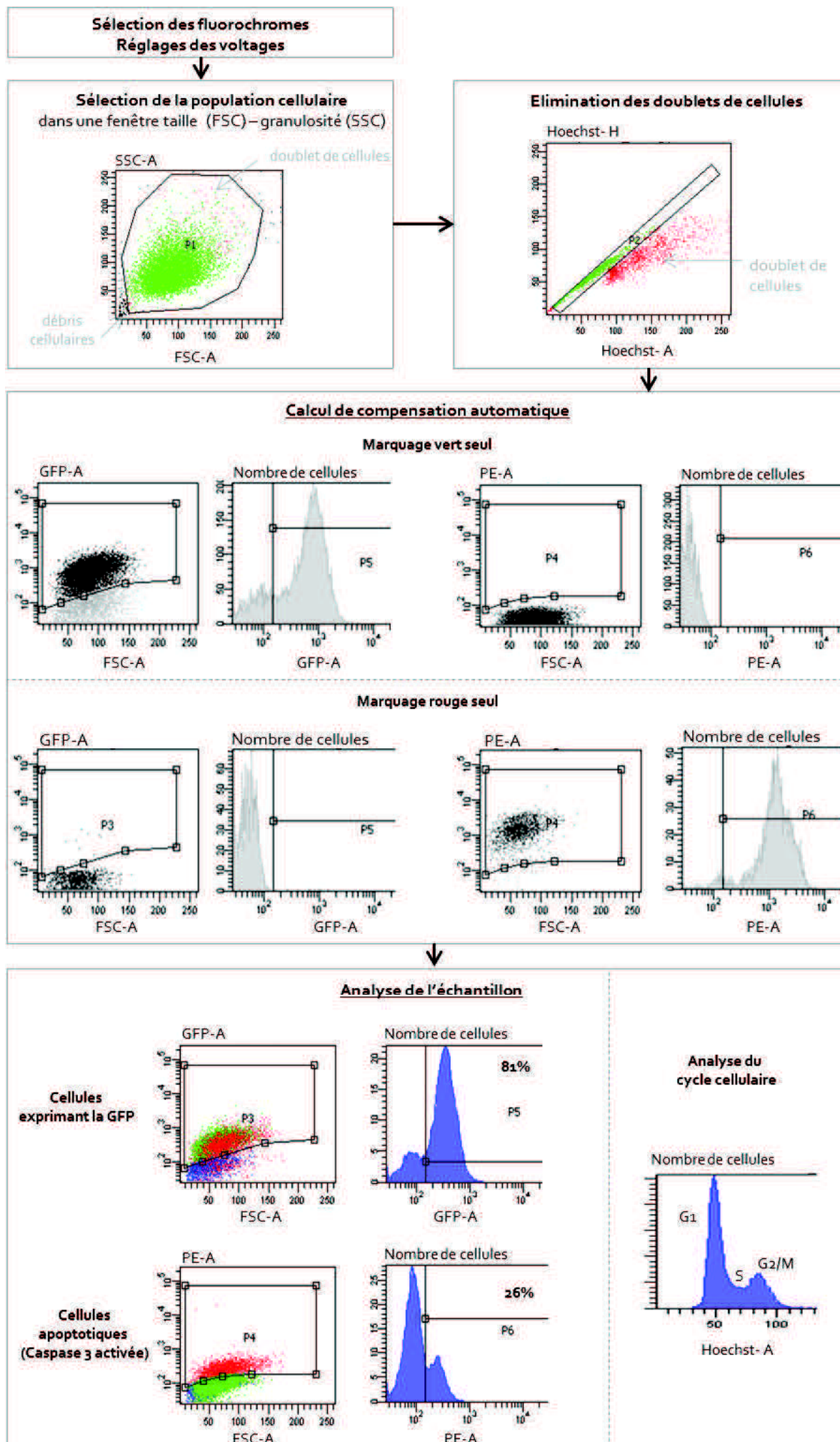


Figure 82 : Aperçu de la méthode d'analyse par cytométrie en flux

sélectionnant que la bissectrice sur le graphe présentant l'aire (Hoechst-A) et la hauteur (Hoechst-H) du signal associé à la fluorescence du Hoechst. Si les deux fluorochromes GFP et PE sont utilisés conjointement, un calcul de compensation permet de supprimer l'effet du recouvrement des spectres d'émission. Le tube contrôle sans marquage permet de définir le seuil d'intensité de fluorescence de PE au-dessus duquel une cellule sera considérée comme marquée. Puis l'intensité de la fluorescence de PE (et éventuellement celle de GFP) est mesurée individuellement pour chaque cellule passant devant le lecteur. Au total, 10 000 événements sont mesurés.

2.1.5.2.2. Tri d'une population cellulaire en cytométrie de flux

Les cellules à trier (exprimant une protéine de fusion GFP) sont décollées par l'action de la trypsine, reprises dans du milieu sans rouge de phénol puis conservées sur glace. Le tri cellulaire est effectué par un cytomètre de flux trieur Aria (BD) grâce au logiciel Diva6 (BD). Le tube contrôle sans marquage permet de définir le seuil d'intensité de fluorescence de GFP au-dessus duquel une cellule sera considérée comme exprimant la protéine de fusion. Après stabilisation du jet, le tri est effectué dans la chambre stérile avec le paramètre « Single Cell », en plaque 96 puits, avec une cellule fluorescente par puits. Une population polyclonale est également triée en mode « Purity » en plaçant 5000 cellules marquées dans un puits de plaque 24 puits. Les clones sont maintenus une semaine en milieu complet sans sélection. Les clones qui se divisent et atteignent la confluence seront passés dans des contenants de plus en plus grands au fil des jours, en maintenant la pression de sélection (300 µg/mL G418).

2.1.5.3. Immunofluorescence sur cellules testiculaires

Les immunomarquages sur cellules spermatogéniques sont réalisés comme décrit dans les précédentes publications de l'équipe (Gaucher et al., 2012 ; Govin et al., 2007). Brièvement, on prélève le testicule d'une souris OF1 et on réalise une congélation-éclair de la gonade dans de l'azote liquide. On coupe le testicule en deux, puis on réalise une empreinte en appliquant la face coupée sur une lame en verre à plusieurs reprises. On fixe les cellules sur la lame par un bain de 10 minutes dans de l'éthanol 90% puis on laisse sécher. On lave 5 minutes dans du PBS. On perméabilise dans 3 bains successifs de 5 minutes dans une solution de perméabilisation Saponine 0,5% / Triton 0,2% / PBS. On lave à nouveau au PBS pendant 5 minutes avant de bloquer durant 1h en PBS/Lait 5% / Tween 0,2%. L'anticorps primaire (**Figure 83**) dilué au 1/50 dans 50 µL de PBS/Tween 0,2% / Lait 1% est placé sur la coupe, entre lame et lamelle pour la nuit, à 4°C et en chambre humide. Après un lavage au PBS pendant 5 minutes, l'anticorps secondaire Alexa 488 (Invitrogen) dilué au 1/500 en PBS/Lait 1% / Tween 0,2% est incubé pendant 30 minutes à 37°C. Après marquage de l'ADN au Hoechst 33258 (Sigma-Aldrich) dilué à 200 ng/mL dans du PBS, les lames sont montées dans du milieu de montage Dakocytomation (Dako) et observées sous microscope à épifluorescence (Olympus).

2.1.5.4. Immunohistochimie sur coupe testiculaire

On réalise des coupes de 6 µm d'épaisseur à l'aide d'un microtome à partir de bloc de paraffine contenant des testicules de souris ou de rat, inclus et colorés à l'éosine. On fait sécher ces coupes sur lame pendant la nuit dans une étuve à 37°C. On déparaffine par 3 bains de 10 minutes de toluène et 2 minutes dans de l'éthanol 100%. On inhibe la peroxydase endogène dans méthanol/0,3% H₂O₂ pendant 30 minutes. On réhydrate par des bains successifs de 2 minutes dans l'éthanol 100%, l'éthanol 90%, l'éthanol 70% puis l'H₂O. On démasque les épitopes par un bain de 20 minutes à 100°C dans une solution de citrate de sodium 10 mM pH=6, on laisse refroidir 20 minutes puis on rince les lames dans H₂O pendant 5 minutes. On bloque les sites non-spécifiques 60 minutes dans du PBS/Lait 5%. On sature 15 minutes avec une solution d'avidine puis de biotine (Avidin/Biotin blocking kit, Vector CliniScience), avec des lavages intermédiaires de 5 minutes dans Lait 0,5% / PBS. L'anticorps primaire (**Figure 83**) dilué au 1/50 dans 50 µL de Lait 1% / PBS est placé sur la coupe, entre lame et lamelle pour la nuit, à 4°C et en chambre humide. On lave deux fois 5 minutes dans un bain de

Antigène	Taille attendue (kDa)	Fournisseur	Espèce	Dilution WB	Dilution IF	Dilution IHC
Actine beta	45	Abcam	souris	1/5000	1/1000	
BRDT C-ter	140 (FL) 90 (ΔC)	SAB 712	lapin	1/2000		
GFP	27	Covance	souris	1/5000		
GM130	130	Abcam	lapin	1/2000	1/200	
GST	25	Millipore	lapin	1/2000		
H3	17	Millipore	lapin	1/5000		
H4 acétylée	15	Upstate 06-598	lapin	1/5000		
HA	1	Covance	souris	1/1000		
HA	1	Roche	rat		1/100	
HDAC1	60	NOTN U309	lapin	1/200		
Lysine acétylée	/	Cell Signaling	lapin	1/5000		
p21 (WAF-1)	21	Oncogene	souris	1/500		
p300	300	Santa Cruz	lapin	1/1000	1/200	
PARP	116 (FL) 89 (clivée)	Boehringer	lapin	1/1000		
PRL	25	Biogenex	souris	1/400		
SOX30	89	Santa Cruz	lapin	1/500	1/50	1/50
SPATA22	41	ProteinTech Group	lapin	1/500	1/50	1/50

Figure 83 : Listes des anticorps utilisés et conditions d'utilisation

Lait 0,5%/PBS. On place 200 μ L de solution d'anticorps secondaire biotinylé (anti-lapin ou anti-souris, Vector) dilué dans Lait 0,5%/PBS sur la lame pour 30 minutes en chambre humide, et on lave à nouveau deux fois 5 minutes dans un bain de Lait 0,5%/PBS. On révèle à l'aide de 200 μ L d'une solution ABC (ABC Elite kit, Vectastain) diluée dans Lait 0,5%/PBS pendant 30 minutes, suivi de deux lavages de 5 minutes dans un bain de Lait 0,5%/PBS. On ajoute 50 μ L d'une solution de 3,3'-diaminobenzidine (DAB) (DAB peroxydase substrate kit ImmPACT, Vector) sur la coupe, on surveille l'apparition du marquage marron sous microscope photonique et on arrête la réaction après 30 secondes à 3 minutes dans un bain d'H₂O pendant 10 minutes. On réalise ensuite une coloration PAS des acrosomes : on place 200 μ L d'acide périodique 37° sur la lame pendant 5 minutes, avant de rincer deux fois 2 minutes à l'eau, et d'ajouter 200 μ L de réactif Schiff 37° pendant 20 minutes. La lame est ensuite rincée à l'eau courante tiède jusqu'à ce que les coupes apparaissent incolores. La contre-coloration est réalisée avec un bain d'hématoxyline de 20 secondes (coloration nucléaire basique), puis d'un bain d'éosine de 15 secondes (coloration cytoplasmique), tous deux suivis d'un rinçage à l'eau courante pendant 1 minute. Les coupes sont enfin déhydratées 2 fois 3 minutes dans l'éthanol 100% puis 2 fois 3 minutes dans le toluène, avant d'être recouvertes d'une goutte de milieu de montage (Eukitt, Fluka) et d'une lamelle qui sera scellée (Rubber Cement). Les lames sont observées à l'immersion sur un microscope photonique (Zeiss) associé à une caméra (AxioCamIC3), et grâce au logiciel AxioVisionLive. Les coupes de tubules séminifères sont stadées en utilisant l'Atlas Russell.

2.2. Méthodes sur acides nucléiques

2.2.1. Extraction d'acides nucléiques

2.2.1.1. Extraction ARN : Kit Macherey-Nagel

Les cellules sont récoltées deux jours après l'ensemencement (et l'éventuelle transfection de siRNA) et après un éventuel traitement, par simple pipettage pour les cellules en suspension, par décollement à la trypsine pour les cellules adhérentes. L'extraction des ARN totaux est réalisée avec le kit « NucleoSpin RNA II » (Macherey-Nagel) selon les recommandations du fournisseur. Brièvement, le culot cellulaire est lysé dans un tampon de lyse contenant du β -mercaptoéthanol. Le lysat est filtré sur colonne pour se débarrasser des gros débris et diminuer la viscosité. L'ARN est précipité puis fixé sur la membrane d'une colonne. Après dessalage de la membrane, les ADN sont digérés directement sur la membrane grâce à l'action d'une DNase recombinante. La membrane est ensuite lavée à deux reprises. Enfin, les ARN sont élués dans un volume de 40 μ L puis dosés au Nanodrop 1000 (Thermo Scientific) après avoir fait le blanc sur le tampon d'élution. Les ratios 260/280 et 260/230 donnent une idée de la contamination par les protéines et les solvants respectivement : ils doivent être proches de la valeur 2. Le ratio 260/280 est souvent bas (compris entre 1 et 1,5) après extraction par ce kit Macherey-Nagel : cela est dû à la forte absorption à la longueur d'onde 280 nm par l'isothiocyanate de guanidium, un sel chaotrope présent dans le tampon de lyse et pas complètement éliminé par les lavages. Les ARN sont conservés à -80°C jusqu'à analyse.

2.2.1.2. Extraction ARN : Kit Qiagen

Le kit RNeasy® Mini Handbook (Qiagen) est utilisé spécifiquement pour les extractions d'ARN en vue d'une analyse transcriptomique pour éviter la contamination à l'isothiocyanate de guanidium. Les cellules sont récoltées deux jours après l'ensemencement et la transfection de siRNA puis centrifugées (3 minutes 3000 rpm). Les cellules sont lysées dans un tampon de lyse puis le lysat est homogénéisé par 5 passages dans une seringue (0,9 mm de diamètre). L'ARN est précipité par l'éthanol 70° puis fixé sur la membrane d'une colonne. La membrane est lavée à trois reprises puis les ARN sont élués dans un volume de 30 μ L de tampon d'élution (en deux passages sur la membrane). Dose et qualité sont mesurés au Nanodrop, puis 2 μ g sont envoyés pour analyse transcriptomique.

Nom	Orientation	Détail de localisation	Séquence
PRL n°1	Forward	Exon 1 (V1), 5'UTR	GCAATCTTGAGGAAGAACTTGA
PRL n°2			CTAGTCTCAGATCTCACCTTC
PRL n°3		Exon 1a (V2), 5'UTR	CTGACGTTTCTATAAAGTAGGTC
PRL n°4			CATTCCAGAAGTACCCTCAAAGAC
PRL n°5			TCGGAACATACAGGCTTTGA
PRL n°6		Exon 1 (V1+V2), 5'UTR	GTGTTTCCTGCAACGATCAC
PRL n°7		Exon 1, CD	ATGAACATCAAAGGATCGCCATG
PRL n°8			AAAGGATCGCCATGGAAAG
PRL n°9		Exon 1 + 2	CATGGAAAGGGTCCCTCCT
PRL n°10		Exon 2	CTCTCCTCAGAAATGTTACGC
PRL n°11		Exon 2+3	GTTACGCGAATTCGATAAACGG
PRL n°12		Exon 3+4	GCCCAACAGATGAATCAAAAAGAC
PRL n°13		Exon 4	CATATTGCGATCCTGGAATGAGC
PRL n°14		Exon 4+5	TGATAGTCAGCCAGGTTTCATC
PRL n°15	Reverse	Exon 2	GCACAGGAGCAGGTTTGAC
PRL n°16		Exon 4	TCCTCAATCTCTACAGCTTTGGA
PRL n°17			GGTTTGCTCCTCAATCTCTAC
PRL n°18		Exon 5, CD	TTAGCAGTTGTTGTTGTGGATG
PRL n°19		Exon 5, 3'UTR	AAGAAGCTTGCAATGGAACG
PRL n°20			CCTGTTACACCCAAGCATGG

Nom	Séquence forward	Séquence reverse
Actine	CATGTACGTTGCTATCCAGGC	CTCCTTAATGTCACGCACGAT
U6	TCGCTTCGGCAGCACA	AACGCTTCACGAATTTGCGT
AUP1	TGCACTCCGTGTACAACAGC	TTTGTGCTTCATGTGCTCT
RELA	CGGGATGGCTTCTATGAGG	CTCCAGGTCCCGCTTCTT
PRL-Rn°1	ACAATGGAGCTTCATGTCCTC	TCCATGAATAGGAGAGTTCTTTAGTTT
PRL-Rn°2	CCGAGAAAGCAGCTGAGTG	CTGTCCTGGATGTAGGCTGAG
BRDT	AAGAAGAGCAAGTTGTGGGTG	TGTACCACGTTCAAGGGAGAA
SOX30	TCCCATCACTCATCCAGTTG	GGCTGGAAAAGTGTGACAGG
SPATA22	TTGGCTCCTGTAATGAAAACAGT	CCATCTCTGTAGCTCCAACCA
p21 = CDKN1A	CTGGAGACTCTCAGGGTCGAA	GGATTAGGGCTTCCTCTTGGA
NBPF4	AGGATGGCAGGATGAAAGAA	AGGACCCCTGCAGTATTGTG
PUMA	GACCTCAACGCACAGTACGA	GAGATTGTACAGGACCCCTCCA
DPEP3	CCTGGACAGGCTTAGAGACG	TCCTGGGACTGGCATGAG
TUBA3C	CGGAGGAGCTCAACATGC	AGTTCACAGCAGGCATTG
RNF17	AGGTCTTCTACCAGCGAATG	GACCGGATGATCTGGATACCC
TRIAP1	GCAGTTCAGGAATTAAGATACTTGG	GCTGGCAATAGCAGATCTTTG
INSIG1	ATCTTTTCCTCCGCTGGT	GGGGTACAGTAGGCCAACAA
PRPF31	GGTGAGCGACTAACGCTAGAA	GAGCTCATCTGCCAGAGACA
RPAIN	TCCCCACACACCTGAATTTT	CACAAGCCCAAGTATCACAGG
MAT2B	TCACATCATTCATGATTTTCAGC	GCATCTGGCTGATTTTCTACAAC
MTHFD2	CAGCTGCAGTTGTGGGAAT	GGAGGCCATCTACATTATCATCA
MT1E	TGCTTGTTCTGCTCACTGGT	GCATTTGCACTCTTTGCACT
PTBP2	GCTCAACAAGCAAACTAGCC	CACAAGTTTGAAAAATCAATCC
PDLIM7	CTGTGCCAAGTGCAAGAAGA	CATGTAGAAGGCCCTGTTCC
RCN2	TGAGAACTGCTCTGGATGAT	TTTTCAAATCGCTTCTGTCC
FERMT2	TTCAGCAACATGAAACAGTGG	TGCAAACTCTACGGTGACCA
RAI14	GATACGCTGAGAGCTCTTCAAAA	TGGTGACTTCCTTCGACATCT

Figure 84 : Liste des primers utilisés

Séquences des principaux primers utilisés en PCR et en qPCR .

2.2.1.3. Extraction ARN au TRizol

Les ARN utilisés pour les analyses en Northern blot sont extraits au TRizol. Environ 10 millions de cellules sont récoltées par centrifugation et lysées dans 1 mL de TRizol (TriReagent, Sigma) pendant 5 minutes. Après ajout de 200 µL de chloroforme, on incube 10 minutes et on centrifuge (15 minutes 12 000 rpm 4°C). On récupère la phase aqueuse supérieure et on précipite les ARN qu'elle contient avec 500 µL d'isopropanol pendant 10 minutes. On isole le culot par centrifugation (10 minutes 14 000 rpm 4°C) et on le lave avec 1,5 mL d'éthanol 70°. Après centrifugation (5 minutes 10 000 rpm 4°C), on sèche le culot pendant 10 minutes et on le resuspend dans 50 µL d'H₂O DEPC. On dose les ARN et on vérifie leur qualité au Nanodrop, puis on utilise 20 µg pour l'analyse en Northern blot.

2.2.2. Analyse des acides nucléiques

2.2.2.1. RT

La réaction de reverse transcription est réalisée grâce au kit RT SuperScript III First-Strand Synthesis SuperMix for qRT-PCR (Invitrogen). Brièvement, 1 µg d'ARN total (extrait ou commercial) est dilué dans 10 µL de tampon de réaction contenant les dNTP, puis on ajoute 1 µL d'enzyme transcriptase inverse, et on complète à un volume final de 20 µL avec de l'H₂O DEPC. On incube 10 minutes à 25°C, 30 minutes à 50°C puis 5 minutes à 85°C. On ajoute 1 µL de RNase H et on place 20 minutes à 37°C. Les ADN complémentaires (ADNc) obtenus sont conservés à -20°C.

2.2.2.2. PCR

Les primers sont désignés grâce à l'un des trois logiciels en ligne suivants : Universal ProbeLibrary Assay Design Center (Roche Applied Science), Primer 3 (<http://frodo.wi.mit.edu/>) ou Primer Bank (<http://pga.mgh.harvard.edu/primerbank/>). Les primers sont désignés pour répondre aux caractéristiques suivantes : température d'hybridation T_m comprise entre 50°C et 60°C (et pas plus de 2°C d'écart pour les primers forward et reverse utilisés conjointement) et taille des primers comprise entre 19 et 25 pb. La spécificité de chaque primer est vérifiée grâce à un BLAST (Basic Local Alignment Search Tool) (<http://blast.ncbi.nlm.nih.gov/Blast.cgi>). Lorsque l'on veut discriminer deux variants ARNm d'un même gène (PRL V1 ou V2), reconnaître plusieurs variants ARNm d'un même gène (6 variants PRL-R) ou discriminer deux membres d'une même famille (NBPF₄, 6 ou 22), on procède à un alignement des séquences nucléotidiques grâce au logiciel en ligne MultiAlin (<http://multalin.toulouse.inra.fr/multalin/>) (Corpet, 1988) et on désigne les primers uniquement sur la partie commune aux différents variants/la partie propre à un variant uniquement. Les primers sont synthétisés par la société Eurogentec à la concentration de 100 mM. La liste des primers utilisés et leurs séquences sont listées dans le tableau ci-contre (**Figure 84**). La réaction de PCR a lieu dans un volume final de 25 µL comprenant 0,5 µL dNTP (1 mM, 5PRIME), 0,5 µL de chacun des primers, forward et reverse (dilué à 10 µM), 2,5 µL de tampon 10X, 0,5 µL d'enzyme TaqPol (5 U/µL, 5PRIME) et les cDNA (produit de RT dilué au 1/10). Le programme de PCR est le suivant : dénaturation à 95°C pendant 1 minute, puis 35 cycles (dénaturation 95°C 30 s, hybridation 30 s, élongation 68°C) et enfin 5 minutes d'élongation terminale à 68°C. Le temps d'élongation à chaque cycle varie en fonction de la longueur du produit de PCR attendu : il est choisi grâce en appliquant le principe « Une minute pour un kilobase ». La température d'hybridation dépend du T_m des primers : souvent 65°, on peut descendre jusqu'à 55°C sans trop d'amplification non-spécifique. Un contrôle négatif est réalisé en remplaçant l'ADN par de l'eau.

2.2.2.3. qPCR

Les analyses en qPCR sont réalisées selon les directives MIQE (Minimum Information for Publication of Quantitative Real-Time PCR Experiments) (Bustin et al., 2009). En qPCR, on utilise des primers forward et reverse pas trop éloignés : la longueur de l'amplicon doit être comprise entre 70 et 200 pb (**Figure 8**). Les

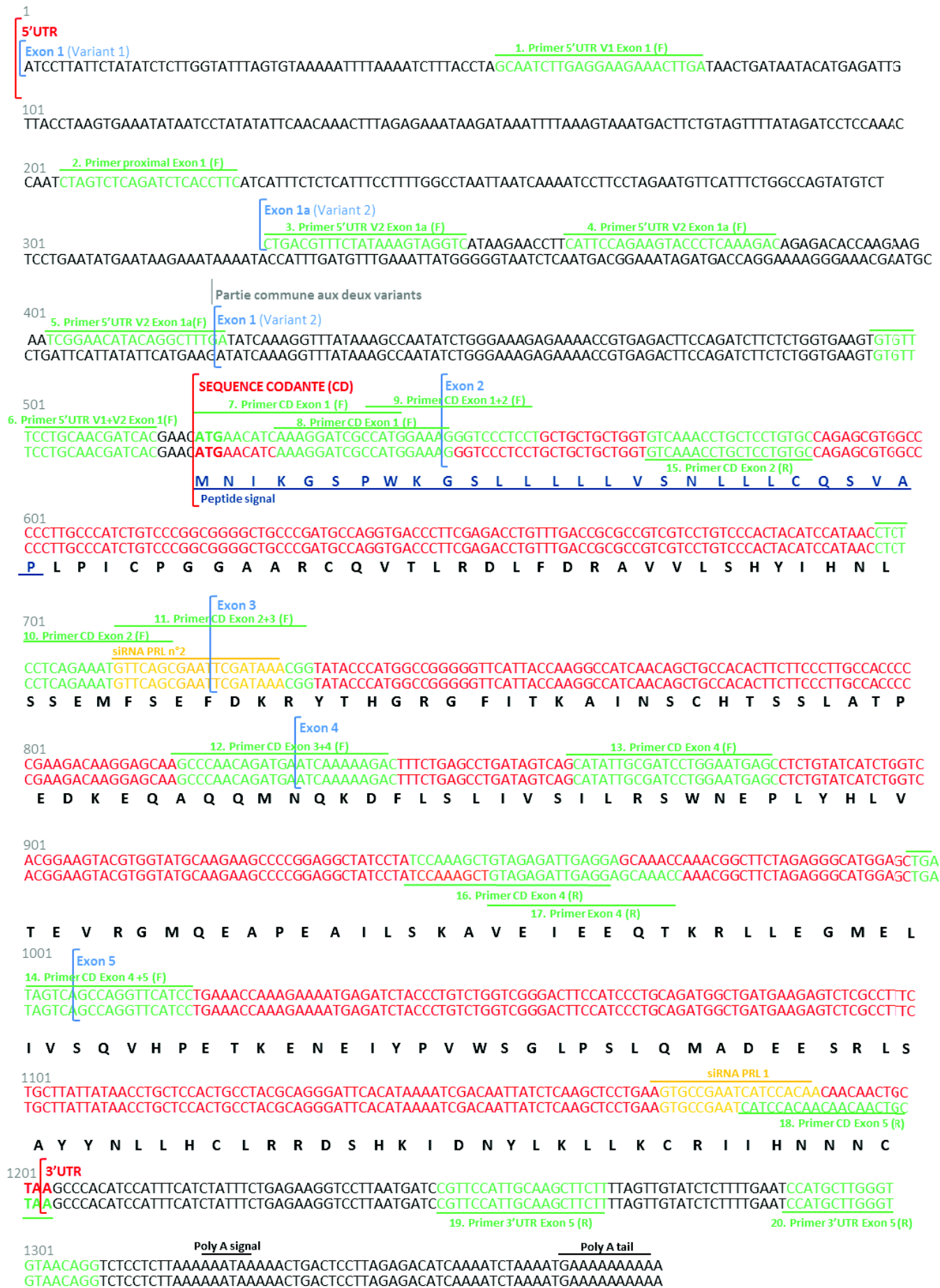
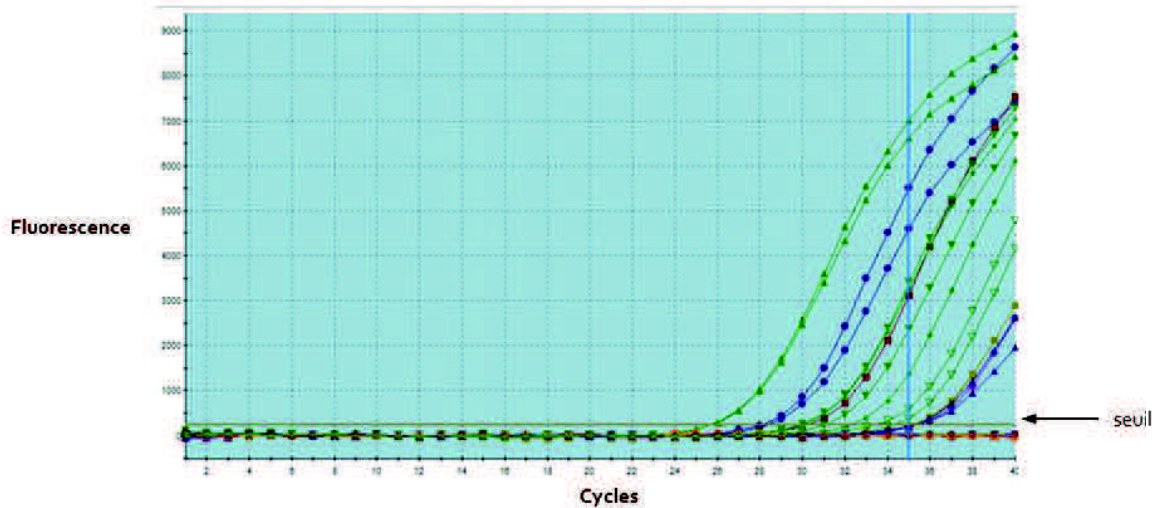
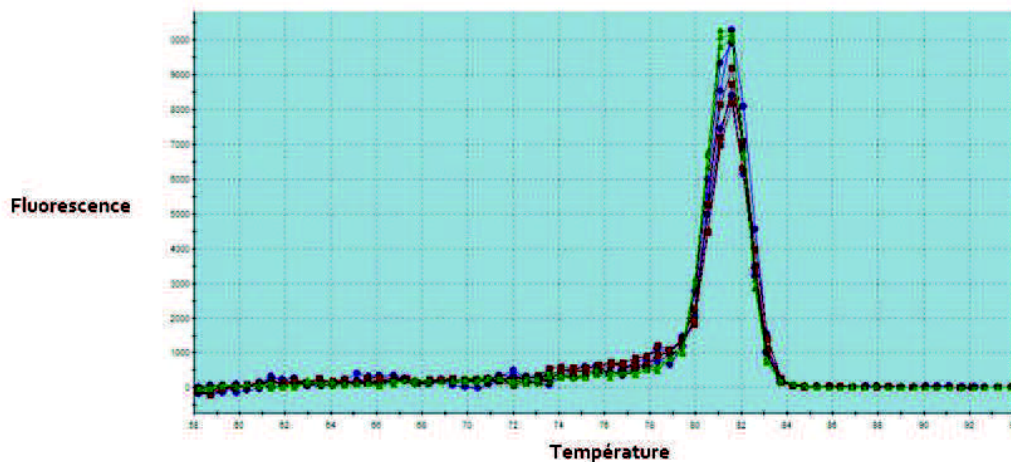


Figure 85 : Localisation des zones d'hybridation des différents primers et siRNA dans la séquence des variants ARNm PRL.

A.



B.



C.

	PRL gene		Control genes	Expression ratio
	Ct	no RT	Mean Ct	PRL/Control genes
Pituitary	15,00	27,97	21,91	706594,43
Placenta	25,74	No Ct	23,04	904,22
Testis	28,26	30,29	21,42	51,29
Uterus	30,44	No Ct	21,11	9,13
Ovary	31,16	No Ct	20,36	3,30
Kidney	33,85	No Ct	21,35	1,01
Liver	No Ct	No Ct	21,84	0,00
Lung	No Ct	No Ct	21,34	0,00
Brain	No Ct	No Ct	21,53	0,00
Bladder	No Ct	No Ct	20,54	0,00
Prostate	No Ct	No Ct	21,23	0,00
Heart	No Ct	No Ct	21,58	0,00
Muscle	No Ct	No Ct	21,41	0,00
Bone marrow	No Ct	No Ct	21,33	0,00
Breast	No Ct	No Ct	20,99	0,00
Colon	No Ct	No Ct	21,00	0,00
Thyroid	No Ct	No Ct	21,39	0,00
B cells	No Ct	No Ct	21,60	0,00
Mean in 6 PRL+ tissues	27,41		21,53	100,00

Figure 86 : Principe de l'analyse des expressions géniques en qPCR

A. Courbes d'amplifications. La valeur de Ct est donnée par le nombre de cycle à partir duquel le signal d'amplification dépasse le seuil.

B. Courbe de dissociation montrant la spécificité des primers utilisés.

C. Exemple d'analyse de données brutes de qPCR, correspondant aux données d'expression du gène *PRL* dans les tissus humains (présentées dans la Figure 34).

Ct = Cycle Threshold

No Ct : Ct > 40

no RT : control on RNA (not submitted to RT)

Dans cet exemple, un ratio d'expression de 100 est affecté à la moyenne de l'expression de *PRL* dans les 6 tissus qui expriment ce gène.

réactions de qPCR on lieu dans les puits d'une plaque 96 puits, dans un volume final de 20 µL. Chaque puits contient 0,03 µL de chaque primer, forward et reverse (100 µM), 10 µL de SYBR Green Brilliant III 2X (Agilent Technology) et 4 µL de cDNA (produit de RT dilué au 1/10). Pour chaque condition, on dépose un duplicat technique et on ajoute un point de contrôle (« no RT ») où le cDNA est remplacé par l'ARN extrait. Cela permet de vérifier l'absence de contamination par l'ADN génomique et le bon fonctionnement de la DNase. La réaction est réalisée dans l'appareil de qPCR Mx3005p cyclé selon le programme suivant (Thermal Profile) : 3 minutes de dénaturation initiale à 95°C, puis 40 cycles comprenant 20 secondes de dénaturation à 95°C et 20 secondes d'hybridation et d'élongation à 60°C.

L'analyse est effectuée grâce au logiciel MxPro (Agilent Technology). On vérifie qu'une paire de primers donnée ne donne bien qu'un seul pic de dissociation. Le seuil est fixé à 235 ; la valeur du Cycle threshold (Ct) est donnée par le nombre de cycles qu'il a fallu atteindre pour que la valeur de fluorescence mesurée dépasse la valeur seuil. Lorsque le Ct dépasse 35, on considère qu'il n'y a pas d'expression. Quand la valeur du Ct obtenu pour le point de contrôle « no RT » est trop proche (noRT-RT<2), on ne prend pas en compte cette valeur. La valeur de Ct analysée est la moyenne des deux valeurs de Ct obtenues pour les deux réplicats techniques (différence < 1). Les valeurs d'expression sont normalisées grâce à des gènes contrôle (*Actin*, *U6*, *RELA* et/ou *AUP1*) choisis car ils présentent les variations minimum parmi les échantillons. La valeur d'expression est obtenue par comparaison avec une condition de référence (hypophyse ou condition contrôle non transfectée par exemple) selon la méthode du « Delta Ct » (Pfaffl, 2001) :

$$\text{Valeur d'expression gène d'intérêt (condition d'intérêt)} = \frac{2^{-\Delta \text{Ct gène d'intérêt (condition de référence - condition d'intérêt)}}}{2^{-\Delta \text{Ct gène contrôle (condition de référence - condition d'intérêt)}}}$$

2.2.2.4. Electrophorèse d'ADN sur gel d'agarose

Le gel d'électrophorèse contient 1% agarose (Invitrogen) et 5 µL de BET (10 mg/mL, Sigma) pour 100 mL de tampon TAE 1X. Les ADN (25 µL de produit de PCR) sont mélangés à 3 µL de bleu de dépôt (bleu de bromophénol, 40% sucrose) et déposés sur le gel, en même temps qu'un marqueur de taille (DNA Ladder, Fermentas ou 1 kb-Plus, Invitrogen). La migration a lieu dans du TAE 1X, à 100 V, pendant environ 30 minutes. La lecture se fait sous lampe à UV.

2.2.2.5. Northern blot

On précipite 20 µg d'ARN total dans de l'acétate de sodium pH=5,2 0,3 M final et 2,5 volumes d'éthanol absolu pendant 30 minutes à -80°C. Après centrifugation (14 000 rpm 15 minutes 4°C), le culot est repris dans 8 µL de tampon d'échantillon (MOPS 1X, formaldéhyde 2,2 M, formamide 50%) puis dénaturé pendant 10 minutes à 60°C avant d'être replacé dans la glace. On ajoute à l'échantillon 4 µL de bleu de dépôt ADN et 2 µL de BET dilué à 0,5 mg/mL. Le gel est composé de 1,2% agarose, MOPS 1X et formaldéhyde complété avec de l'H₂O DEPC. Après dépôt des échantillons sur gel, la migration est lancée lentement (20 V) dans du tampon de migration (MOPS 1X, formaldéhyde 2,2 M) et se fait sur la nuit. Après observation aux UV, le gel est équilibré dans du tampon SSC 20 X (Euromedex), puis les ARN sont transférés sur membrane de nitrocellulose (Amersham Hybond-N, GE Healthcare) préalablement équilibrée dans l'eau. Le transfert se fait sur la nuit par capillarité, grâce à un pont salin imbibé de tampon SSC 20X. La membrane est rincée dans du tampon SSC 3X puis les ARN sont fixés 2h à 80°C sous vide. On pré-hybride la membrane, après réhydratation dans H₂O, pendant 2h dans une rôtissoire à 65°C dans 20 mL de tampon d'hybridation (Na₂HPO₄ pH=7,2 0,25 M, SDS 7%) préalablement chauffé. Pendant ce temps, on synthétise la sonde marquée. Le produit PCR (taille proche de 300 pb) a préalablement été purifié après électrophorèse sur gel d'agarose (NucleoSpin Gel and PCR Clean-Up, Macherey Nagel) et dosé au Nanodrop. Le marquage de la sonde se réalise par la technique du Random Priming (Random Primed DNA Labeling Kit, Roche) selon les recommandations du fournisseur. On dilue 2 µL du produit PCR (12,5 ng/µL) dans 5 µL d'eau, on dénature 5 minutes dans de l'eau bouillante et on

place immédiatement sur glace. On ajoute 2 µL de tampon contenant les hexanucléotides, 3 µL d'un mix de désoxynucléotides (dATP, dGTP, dTTP), 5 µL (20 µCi) de dCTP marqué au α - ^{32}P et enfin 1 µL d'enzyme de Klenow. On incube 30 minutes à 37°C, et on stoppe la réaction en ajoutant 2 µL EDTA 0,2M pH=8 pendant 10 minutes à 65°C. Puis on purifie sur colonne après avoir complété avec du tampon à un volume final de 50 µL (ProbeQuant, GE Healthcare). On ajoute à la sonde marquée 750 µL d'ADN de sperme de saumon (DNA fish sperm MB grade, 10 mg/mL, Roche), on dénature le tout 10 minutes dans de l'eau bouillante et on replace 5 minutes sur glace. On ajoute ce mélange à 15 mL de tampon d'hybridation préalablement préchauffé et on incube avec la membrane sur la nuit dans la rôtissoire à 65°C. La membrane est lavée puis séchée. On expose un film (Kodak) pendant 5 jours à -80°C et on révèle en chambre noire

2.2.3. Clonage

Les plasmides GFP-PRL, PRL-GFP, HA-PRL, GFP-PRL16, 6His-PRL et GST-PRL ont été créés pour les besoins de cette étude. Les primers utilisés pour l'amplification de l'insert sont désignés pour encadrer la séquence à cloner (PRL ou PRL16) contenant ou non le codon stop et pour respecter le cadre ouvert de lecture permettant la production d'une protéine de fusion. Les primers incluent des sites de restriction (BamHI, KpnI, NotI, NheI). La présence d'un site EcoRI dans la séquence codante du gène PRL est révélée par l'analyse en ligne avec l'outil NEBcutter V2.0 (<http://tools.neb.com/NEBcutter2/>) : l'enzyme EcoRI ne sera donc pas utilisée. La séquence codant la prolactine est amplifiée par PCR à partir d'ADNc hypophysaire grâce au kit haute fidélité Advantage II PCR Kit (Clontech). L'amplicon est digéré pendant 2 heures à 37°C dans un volume final de 50 µL, avec les deux enzymes de restriction (BioLabs) et le tampon approprié. Environ 3 µg de vecteur (peGFP-C3, peGFP-N2, pcDNA3.1/HA-C, pET-28a+ ou pGEX-5X-1) sont digérés en parallèle par les mêmes enzymes dans les mêmes conditions. Le vecteur digéré est ensuite déphosphorylé par une incubation de 30 minutes à 37°C avec la CIP pour éviter sa re-ligation sur lui-même. L'insert et le vecteur digéré sont déposés sur gel, puis purifiés après découpage des bandes grâce au kit NucleoSpin Extract II (Macherey-Nagel) selon les recommandations du fournisseur. Un nouveau dépôt de 1 µL d'insert et de vecteur purifiés permet la quantification avant la ligation. La ligation est réalisée dans un volume de 10 µL, avec un ratio insert : vecteur de 3 et en présence de T4 DNA ligase (BioLabs) pendant la nuit à 16°C.

On transforme 30 µL de bactéries compétentes BL21 Gold (Agilent) avec les 20 µL de produit de ligation par choc thermique (90 secondes à 42°C immédiatement suivies de 90 secondes sur glace). On laisse les bactéries transformées pousser pendant 1 heure en milieu LB liquide à 37°C avant de les étaler sur du LB Agar additionné de l'antibiotique approprié (Ampicilline ou Kanamycine, Sigma) pour la nuit à 37°C. On repique une dizaine de colonies et on lance des cultures de 5 mL en LB + Antibiotique pendant 2h à 37°C sous agitation. On purifie alors le plasmide de chacun de ces clones grâce au kit de Mini Prep NucleoSpin Plasmid (Macherey-Nagel) selon les recommandations du fournisseur. Les différents clones sont vérifiés par restriction en digérant 5 µL des plasmides purifiés par les enzymes utilisées pour le clonage dans un volume final de 20 µL et en déposant 5 µL des produits de la double digestion sur gel d'agarose. Les plasmides présentant un profil de restriction correct sont envoyés pour séquençage à la société Cogenics. Le plasmide obtenu est conservé à -20°C et utilisé pour les transfections de cellules humaines en culture (**voir paragraphe 2.1.3.2**) ou transformation de bactéries compétentes pour la production de protéines recombinante (**voir paragraphe 2.3.2**).

2.3. Méthodes sur protéines

2.3.1. Extraction des protéines

2.3.1.1. En tampon de lyse dénaturant

Les cellules sont récoltées, centrifugées (3 minutes 3000 rpm) et le culot est lavé avec 1 mL PBS. Le culot est ensuite repris dans 200 µL de tampon de lyse dénaturant (Tris HCl 0,5 M pH=6,8, NaCl 500 mM, Glycérol 20%, NP40 1%, SDS 0,1%, bêta-mercaptoéthanol 1%), en présence d'anti-protéases (Complete, Roche). Le lysat visqueux est soniqué à 200 J sur glace puis incubé 30 minutes sur glace. Une centrifugation (14 000 rpm, 10 minutes, 4°C) permet de se débarrasser des débris cellulaires. Le contenu protéique du surnageant est dosé par la méthode Bradford avec mesure de l'absorbance à une longueur d'onde de 595 nm et calcul de la concentration grâce à une gamme d'albumine sérique bovine BSA (0 – 8 µg/L). L'extrait protéique est conservé à -20°C.

2.3.1.2. En tampon de lyse douce

Les cellules sont récoltées, centrifugées (3 minutes 3000 rpm) et le culot est lavé avec 1 mL PBS. Le culot est ensuite repris dans 100 µL de tampon LSDB (Glycérol 20%, MgCl₂ 3 mM, HEPES 50 mM, NP40 1% et antiprotéases) additionné de KCl 500 mM ou de NaCl 500 mM. Le lysat est incubé 30 minutes sur glace puis centrifugé (14 000 rpm, 10 minutes, 4°C) ; le surnageant peut être dosé selon la méthode Bradford et est utilisé pour les expériences de *déphosphorylation* in vitro, d'immunoprécipitation (IP), de test d'interaction in vitro avec protéine recombinante ou de peptide pull-down (PPD).

2.3.1.3. Fractionnement protéique cytosoluble-nucléaire insoluble

Les cellules sont récoltées, centrifugées (3 minutes 3000 rpm) et le culot est lavé avec 1 mL PBS. Le culot est ensuite repris dans 200 µL PBS/Triton 0,05% additionné d'antiprotéases puis incubé 10 minutes sur glace. Après centrifugation (5 minutes 2500 rpm 4°C), le surnageant correspond à la fraction cytosoluble. Le culot de noyaux est lavé dans 200 µL PBS, centrifugé (5 minutes 2500 rpm 4°C), repris dans 200 µL de tampon de lyse dénaturant (**voir paragraphe 2.3.1.1**) puis soniqué à 200 J sur glace. Après 30 minutes d'incubation sur glace et centrifugation (10 minutes 14 000 rpm 4°C), le surnageant correspond à la fraction nucléaire. Les extraits protéiques cytosoluble et nucléaire sont conservés à -20°C. Un volume de 10 µL d'extrait, dilué avec 10 µL d'H₂O, sera utilisé pour électrophorèse.

2.3.1.4. Extrait acide des histones

Les cellules sont récoltées, centrifugées (3 minutes 3000 rpm) et le culot est lavé avec 1 mL PBS. Le culot est ensuite repris dans 200 µL PBS/Triton 0,05% additionné d'antiprotéases puis incubé 10 minutes sur glace. Après centrifugation (5 minutes 2500 rpm 4°C), le culot de noyaux est lavé dans 200 µL PBS puis centrifugé (5 minutes 2500 rpm 4°C). L'extraction acide se fait par incubation des noyaux sur la nuit, à 4°C et sous agitation dans 50 µL d'HCl 0,2 N. Après centrifugation (10 minutes 14 000 rpm 4°C), le surnageant correspond à l'extrait acide enrichi en protéines nucléaires basiques, histones principalement. Ces protéines peuvent être ensuite précipitées dans TCA 18% 30 minutes à 4°C sous agitation. Le précipité est culoté par centrifugation (10 minutes 10 000 rpm 4°C), lavé deux fois dans 500 µL d'acétone, séché, puis repris dans 100 µL H₂O avec antiprotéases. L'extrait est conservé à -20°C. Un volume de 5 µL d'extrait, dilué avec 15 µL d'H₂O, sera utilisé pour électrophorèse.

2.3.2. Production de protéine recombinante

2.3.2.1. Cinétique d'induction

Un clone bactérien BL21 Gold (Agilent) transformé par le plasmide d'intérêt, et un clone contrôle transformé par le plasmide vide sont mis en pré-culture à 37°C pendant 2h sous agitation dans 5 mL de LB additionné de l'antibiotique de sélection approprié (Kanamycine ou Ampicilline). Puis on lance une culture de 10 mL en inoculant avec 1 mL de préculture. Lorsque la mesure d'absorbance à $\lambda=600$ nm atteint une densité optique $DO=0,4$ et après avoir prélevé 1 mL de culture non induite, on induit la production de la protéine recombinante d'intérêt avec 0,5 mM IPTG (stock 200 mM). On réalise une cinétique d'induction en prélevant 1 mL de culture induite après 15, 30, 45 minutes, 1h, 1h30, 2h et 3h. Des extraits protéiques totaux sont préparés sur les différents points prélevés avant et après induction : les bactéries sont centrifugées (30 secondes 11 000 rpm) et lavées dans 1 mL de PBS. Le culot est lysé dans 50 μ L de bleu de dépôt de protéines (SDS) 2X par sonication à 200 J. Après centrifugation (10 minutes 14 000 rpm 4°C), on récupère le surnageant qui peut être dénaturé et déposé sur gel d'électrophorèse. On observera en Coomassie la cinétique d'induction de la production de la protéine recombinante.

2.3.2.2. Extrait protéique soluble

Quand la durée d'induction adéquate est déterminée, on lance une culture dans 20 mL LB additionné de l'antibiotique à 37°C sous agitation et on induit avec 0,5 mM d'IPTG quand $DO=0,4$. A l'issue de la durée d'induction, on réalise un extrait protéique soluble par lyse douce. Les bactéries sont centrifugées (15 minutes 4000 rpm 4°C), lavées avec 1 mL PBS et lysées dans 1 mL de tampon LSDB 500 mM NaCl en soniquant à 200 J. Après 30 minutes d'incubation sur glace, les débris sont culotés (10 minutes 14000 rpm 4°C) et le surnageant correspond à l'extrait soluble. L'extrait protéique insoluble peut être obtenu en reprenant le culot dans 200 μ L de bleu de dépôt de protéines (SDS) 2X, en soniquant à 200 J, en centrifugeant (10 minutes 14000 rpm 4°C) et en conservant le surnageant. Si la protéine recombinante d'intérêt se révèle après électrophorèse être produite sous forme insoluble, on diminue la quantité d'IPTG (jusqu'à 0,1 mM) e/ou on induit sur la nuit à température ambiante et non plus à 37°C.

2.3.2.3. Purification d'une protéine recombinante tagguée 6-histidines sur billes agarose-nickel

On prélève 50 μ L de billes Agarose-Nickel que l'on centrifuge (2 minutes 2000 rpm) pour supprimer le tampon et que l'on équilibre par deux lavages de 5 minutes à 4°C sous agitation dans 1 mL de tampon LSDB 500 NaCl additionné de 20 μ M d'imidazole. On utilise 50 μ L d'extrait soluble induit comme input, et 500 μ L de ce même extrait additionné de 20 μ M d'imidazole que l'on incube avec les billes pendant 1h à 4°C sous agitation. On lave les billes 3 fois 5 minutes avec 1 mL de tampon à 4°C sous agitation. Si l'on souhaite vérifier l'efficacité de la purification, on élue les billes avec 50 μ L de bleu de dépôt de protéines (SDS) 2X pendant 5 minutes à 95°C, on centrifuge (3 minutes 14000 rpm) et on analyse le surnageant par électrophorèse. On peut également éluer la protéine purifiée dans 50 μ L de tampon LSDB 500 mM NaCl additionné de 250 μ M imidazole pendant 30 minutes à 4°C sous agitation. Les billes portant la protéine purifiée peuvent également être utilisées directement pour des tests d'interaction in vitro. Un contrôle négatif est réalisé avec un extrait protéique de bactéries non transformées ou une culture bactérienne non induite.

2.3.2.4. Purification d'une protéine recombinante fusionnée à la GST sur billes sépharose glutathion

On prélève 50 μ L de billes Sépharose-GST que l'on centrifuge (5 minutes 5000 rpm) pour supprimer le tampon et que l'on lave deux fois en inversant dans 100 μ L de PBS. On utilise 50 μ L d'extrait soluble induit comme input, et 500 μ L de ce même extrait que l'on incube avec les billes pendant 30 minutes à température

ambiante sous agitation. On lave les billes 3 fois dans 100 µL de PBS. Si l'on souhaite vérifier l'efficacité de la purification, on élue les billes avec 25 µL de bleu de dépôt de protéines (SDS) 2X pendant 5 minutes à 95°C, on centrifuge (3 minutes 14000 rpm) et on analyse le surnageant par électrophorèse. On peut également éluer la protéine purifiée dans 25 µL de Tris-HCl 50 mM à pH=8 additionné de 10 mM de glutathion réduit pendant 10 minutes à température ambiante sous agitation. Les billes portant la protéine purifiée peuvent également être utilisées directement pour des tests d'interaction *in vitro*. Un contrôle négatif est réalisé avec un extrait protéique de bactéries non transformées ou une culture bactérienne non induite ; un contrôle positif est réalisé en purifiant la GST seule à partir d'un extrait protéique soluble de bactéries transformées par le plasmide pGEX vide.

2.3.3. Tests *in vitro* sur protéines

2.3.3.1. Expérience de déphosphorylation *in vitro* des protéines

Les protéines totales, extraites par lyse douce en tampon LSDB 500, sont soumises à déphosphorylation par la phosphatase intestinale de veau (calf intestine phosphatase CIP). 50 µg de protéines sont dilués dans 100 µL de tampon CIP (Tris-HCl 50 mM, MgCl₂ 10 mM, NaCl 100 mM, DTT 1 mM avec antiprotéases). Elles sont incubées 30 minutes à 37°C avec 50 unités de CIP (10 000 units/mL). Un tube contrôle est incubé en absence d'enzyme. Environ 10 µg de protéines seront utilisées pour électrophorèse.

2.3.3.2. Immunoprécipitation

On dispose d'un extrait protéique de cellules eucaryotes (éventuellement transfectées) en tampon de lyse douce LSDB 500 mM KCl ; on en conserve 10 µL comme input. On abaisse la concentration des 90 µL restants à 250 mM KCl. On incube sur la nuit à 4°C et sous agitation avec 2 µg d'anticorps. Un contrôle négatif est réalisé en parallèle avec un anticorps non pertinent. On prépare 20 µg de Dynabeads protéine G en les lavant 2 fois 5 minutes avec 500 µL de tampon LSDB 250 mM à 4°C sous agitation. On sature les billes 30 minutes à 4°C sous agitation avec 1 mL de tampon LSDB 250 mM KCl additionné de 500 µg de BSA. On ajoute sur billes l'extrait protéique incubé avec l'anticorps et on agite 1h à 4°C. On lave les billes 2 fois 5 minutes avec 500 µL de tampon LSDB 250 mM à 4°C sous agitation, puis on dessale avec 500 µL de PBS pendant 5 minutes à 4°C sous agitation. On élue alors en reprenant les billes dans du bleu de dépôt pour protéines (SDS) 2X pendant 5 minutes à 95°C. On ne récupère que le surnageant qui sera déposé tel quel sur le gel d'électrophorèse.

2.3.3.3. Test d'interaction *in vitro* avec protéine purifiée sur billes

D'une part, on dispose d'une protéine recombinante purifiée sur billes, billes que l'on équilibre trois fois 5 minutes dans 500 µL de tampon LSDB 250 mM NaCl à 4°C sous agitation. D'autre part, on dispose d'un extrait protéique de cellules eucaryotes (éventuellement transfectées) en tampon de lyse douce LSDB 500 mM NaCl, dont on abaisse la concentration en sels jusqu'à 250 mM. On incube l'extrait cellulaire avec les billes pendant 1h à 4°C sous agitation. On lave les billes deux fois 5 minutes dans 500 µL de tampon LSDB 250 mM NaCl puis on dessale avec 500 µL de PBS additionné d'antiprotéases pendant 5 minutes à 4°C sous agitation. On élue en reprenant les billes dans 20 µL de bleu de dépôt de protéines (SDS) 2X pendant 5 minutes à 95°C. Le surnageant est analysé par électrophorèse. Un contrôle négatif est réalisé en incubant l'extrait cellulaire avec les billes seules ou avec les billes mise au contact de l'extrait bactérien non induit par l'IPTG.

Gel de séparation						Gel de concentration
	6% > 150 kDa p300	8% 80-150 kDa BRDT	10% 40-80 kDa SOX30, SPATA22, PRL-GFP	12% 20-40 kDa PRL	16,5% < 20 kDa Histones	
Tris 1,5 M pH=8,8	1,3 mL					/
Tris 1 M pH=6,8	/					250 µL
Acrylamide/ Bisacrylamide 30%	1 mL	1,3 mL	1,7 mL	2 mL	2,5 mL	330 µL
SDS 10%	50 µL					20 µL
APS	50 µL					20 µL
TEMED	4 µL					2 µL
Volume final	5 mL					2 mL

Figure 88 : Composition des gels de concentration utilisés pour électrophorèse sur gel d'acrylamide en fonction de la taille de la protéine d'intérêt

2.3.3.1. Peptide pull-down

On réalise tout d'abord le couplage des billes et des peptides. On prélève 50 µL de billes magnétiques recouvertes de streptavidine Dynabeads Streptavidin. On élimine le tampon en plaçant les tubes sur un portoir magnétique. On lave les billes avec 1 mL de tampon LSDB 500 mM NaCl puis 1 mL de PBS, additionnés d'antiprotéases pendant 1 minute à 4°C sous agitation. On place sur les billes le peptide (peptide H₄ acétylé) dilué à 5 µM dans 1 mL de PBS et on incube sur la nuit à 4°C sous agitation. On lave les billes avec 1 mL de tampon LSDB 500 mM NaCl pendant 1 minute à 4°C sous agitation. On dispose d'un extrait protéique de cellules eucaryotes (éventuellement transfectées) en tampon de lyse douce LSDB 500 mM NaCl. On incube l'extrait cellulaire avec les billes pendant 2h à 4°C sous agitation. On lave les billes deux fois 5 minutes dans 1 mL de tampon LSDB 500 mM NaCl. On élue en reprenant les billes dans 20 µL de bleu de dépôt de protéines (SDS) 2X pendant 5 minutes à 95°C. Le surnageant est analysé par électrophorèse. Les contrôles négatifs sont réalisés en utilisant un peptide contrôle (peptide H₄ non acétylé) et un extrait cellulaire de cellules non transfectées.

2.3.4. Western-blot

2.3.4.1. Electrophorèse SDS-Page

Différentes concentrations en acrylamide du gel de séparation sont utilisées selon la taille de la protéine d'intérêt (**Figure 88**). Le gel de séparation est coulé dans le système de montage (épaisseur 1 mm) et polymérise en étant isolé de l'air par une pellicule d'isopropanol. Le gel de concentration est ensuite coulé au-dessus, et on ajoute le peigne qui formera les puits de dépôt. Entre 1 et 20 (Western blot) ou 10 et 50 µg (Coomassie) de protéines sont alors déposées sur le gel. L'extrait protéique est dilué à la concentration voulue dans H₂O avant dépôt, est additionné de bleu de dépôt de protéines (SDS 5X) et est dénaturé 5 minutes à 95°C. On dépose également 5 µL de marqueur de poids moléculaires (Prestained Protein Ladder, Fremontas). La migration a lieu dans du tampon TGS 1X (0,75g/L Tris, 3,6 g/L Glycine, 0,1% SDS) pendant au moins 1h à 180 V. Le gel est démoulé et peut être coloré par le bleu de Coomassie pendant 2h sous agitation, décoloré sur la nuit sous agitation dans du tampon de décoloration, réhydraté 1h dans H₂O et séché.

2.3.4.2. Transfert

Les protéines du gel peuvent être transférées sur une membrane de nitrocellulose (Amersham Hybond-C, GE Healthcare) dans du tampon de transfert à l'éthanol (TGS 50%, éthanol 20%) ou à l'isopropanol pour le transfert de p300 (TGS 50%, isopropanol 10%) pendant 1h à 100V (ou 1h30 à 110V pour les protéines de grande taille comme BRDT et p300). Le bon déroulement du transfert est vérifié en colorant les protéines sur la membrane au rouge ponceau (Sigma).

2.3.4.3. Immunodétection

La membrane est saturée 1h sous agitation dans du PBS/Lait 5%/Tween 0,1%. Elle est incubée sur la nuit à 4°C sous agitation avec l'anticorps primaire dilué à la concentration adéquate (**Figure 83**) dans du PBS/Lait 1%/Tween 0,1%. La membrane est lavée 3 fois 10 minutes sous agitation dans du PBS/Tween 0,1% puis incubée pendant 1h sous agitation avec l'anticorps secondaire couplé à la peroxidase de raifort HRP (anti-souris ou anti-lapin IgG, Dako) dilué au 1/10 000 dans du PBS/Lait 1%/Tween 0,1%. Pour les expériences d'immunoprécipitation, on utilise des anticorps secondaires anti-chaînes légères ou anti-chaînes lourdes pour ne pas être gêné par le signal des IgG. Après 3 derniers lavages de 10 minutes sous agitation dans du PBS/Tween 0,1%, on procède à la révélation avec le kit ECL Plus (GE Healthcare) selon les recommandations du fournisseur. Après incubation de la membrane dans la solution de substrat, on expose la membrane à un film photosensible (Kodak) en chambre noire. Après exposition, le signal est révélé par des bains successifs dans des solutions de révélation et de fixation. Pour un nouveau western-blot, la membrane peut être débarrassée des anticorps par incubation dans une solution de stripping (SDS 2%, bêta-mercaptoéthanol 0,7%, Tris-HCl 50 mM pH=6,8)

pendant 20 minutes à 50°C, suivie de nombreux lavages au PBS/Tween 0,1% et d'un nouveau blocage en PBS/Lait 5%. Les signaux obtenus peuvent être quantifiés grâce au logiciel ImageJ (Image Processing and Analysis in Java).

2.4. Méthodes sur données bioinformatiques

2.4.1. Patients

Les données transcriptomiques issues des tumeurs pulmonaires de la cohorte « Cancer du poumon » ont été obtenues avec la technologie Affymetrix ("Affymetrix.GeneChip.HG-U133_Plus_2"). Ces données, dont l'obtention a été précédemment décrite (Rousseaux, 2013) sont disponibles sur le site GEO sous le référencement GSE30219. L'analyse des échantillons cancéreux a été menée conjointement sur 112 échantillons de tissus humains, incluant par exemple des échantillons d'ovaire, de testicule, de sein ou d'hypophyse.

Les analyses d'expression des gènes *PRL* et *PRL-R* dans les échantillons de patients ont été réalisées en comparant les signaux fournis par les sondes 205455_at (*PRL*) et 205455_at (*PRL-R*) à des seuils que nous avons établis. Pour le gène *PRL*, le seuil au-dessus duquel on considère que le gène est exprimé correspond à la moyenne des signaux de la sonde 205455_at dans les 109 échantillons de tissus humains, sans prendre en compte les 3 échantillons hypophysaires. Pour le gène *PRL-R*, le seuil au-dessus duquel on considère que le gène est exprimé correspond à la moyenne des signaux de la sonde 205455_at dans les 112 échantillons à laquelle on ajoute deux fois la valeur de l'écart-type (Mean + 2 standard deviation (sd)).

Les analyses de survie ont été réalisées par Alexandra Debernardi selon la méthode de Kaplan-Meier (Kaplan, 1958) dans une analyse univariée à l'aide du logiciel R (version 2.10.0). Les probabilités de survie sont comparées entre les patients dont la tumeur exprime le gène (signal au-dessus du seuil) et les autres. Un test du logrank Mantel-Cox permet de comparer les deux courbes de survie. Les analyses de pronostic sont validées en appliquant la même méthode sur l'étude externe GSE1988 disponible en ligne sur le portail GEO (<http://www.ncbi.nlm.nih.gov/geo/>). Pour évaluer le pouvoir discriminant de l'expression du gène par rapport aux autres paramètres influant sur le pronostic de la tumeur pulmonaire, on réalise une analyse multivariée selon le modèle de Cox prenant en compte les types histologiques et les paramètres TNM.

Les comparaisons des transcriptomes des patients *PRL+* et *PRL-* ont été réalisées grâce au logiciel Gene Spring. Les gènes exhibant des différences significatives d'expression entre les patients *PRL+* et les patients *PRL-* sont identifiés par un test de Mann-Whitney avec un seuil $p < 0,01$ et un fold change à 1,5. La heatmap est obtenue avec le logiciel PermutMatrix (version 1.9.3 disponible en ligne (<http://www.lirmm.fr/~caraux/PermutMatrix/>)) en représentant les valeurs RMA normalisées selon une échelle de couleur (vert : faible expression, rouge : expression forte). Les fonctions enrichies dans chaque liste sont identifiées en utilisant les termes GO. La valeur - log corrected p-value est représentée comme un indicateur d'enrichissement. Les mesures de similarité pour identifier des gènes corégulés avec *PRL* sont réalisées avec les tests de Spearman et de Pearson avec des seuils à 0,06 grâce au logiciel Genespring.

2.4.2. Analyse transcriptomique des lignées CBPC

Six répliquats de cellules de CBPC H146 et H524 ont été transfectées conjointement avec les deux couples de siARN *PRL* (voir paragraphe 2.1.3.1). Après 48h, les cellules sont récoltées et les ARN totaux sont extraits (voir paragraphe 2.2.1.2). L'efficacité du silençage dans les 6 répliquats a été vérifiée par RT-qPCR avant envoi des échantillons (Figure 89). L'analyse transcriptomique a été réalisée par la Plateforme SNP, Transcriptome & Epigénomique de l'hôpital d'Angers dirigée par le Docteur Guardiola. Les données sont normalisées par les quantiles. Les répliquats n°6 (Figure 86, Manip F) sont exclus de l'analyse qui se poursuit sur cinq répliquats

biologiques. Pour chaque sonde, on calcule la moyenne d'expression dans chaque condition (si – ou siPRL) en moyennant le signal des 5 réplicats. On réalise ensuite pour chaque réplicat un test de distribution : ne sont conservées que les sondes dont les valeurs de signal de 4 réplicats au moins sont comprises dans l'intervalle moyenne ± 2 sd. Parmi ces sondes, on considère comme différentiellement exprimées celles qui présentent un fold change supérieur à 1,2.

L'analyse de Gene Ontology est réalisée avec le logiciel AMEN ([Chalmel and Primig, 2008](#)) et ne présente que les termes significativement enrichis. L'analyse GSEA (Gene Set Enrichment Analysis) est une méthode développée pour interpréter les résultats d'analyse différentielle à l'aide de différentes sources d'information biologique. Elle permet de mesurer le recouvrement de notre liste de gènes avec des groupes de gènes différentiellement exprimés entre deux conditions biologiques données, répertoriés dans la base de données du Broad Institute (<http://www.broadinstitute.org/gsea/>) ([Tamayo et al. ,2005](#)). Enfin, la conversion des sondes Illumina en identifiants Affymetrix, pour comparaison des données obtenues en cellules et en patients, est réalisée grâce au portail Biomart Central Portal disponible en ligne (<http://central.biomart.org/>).

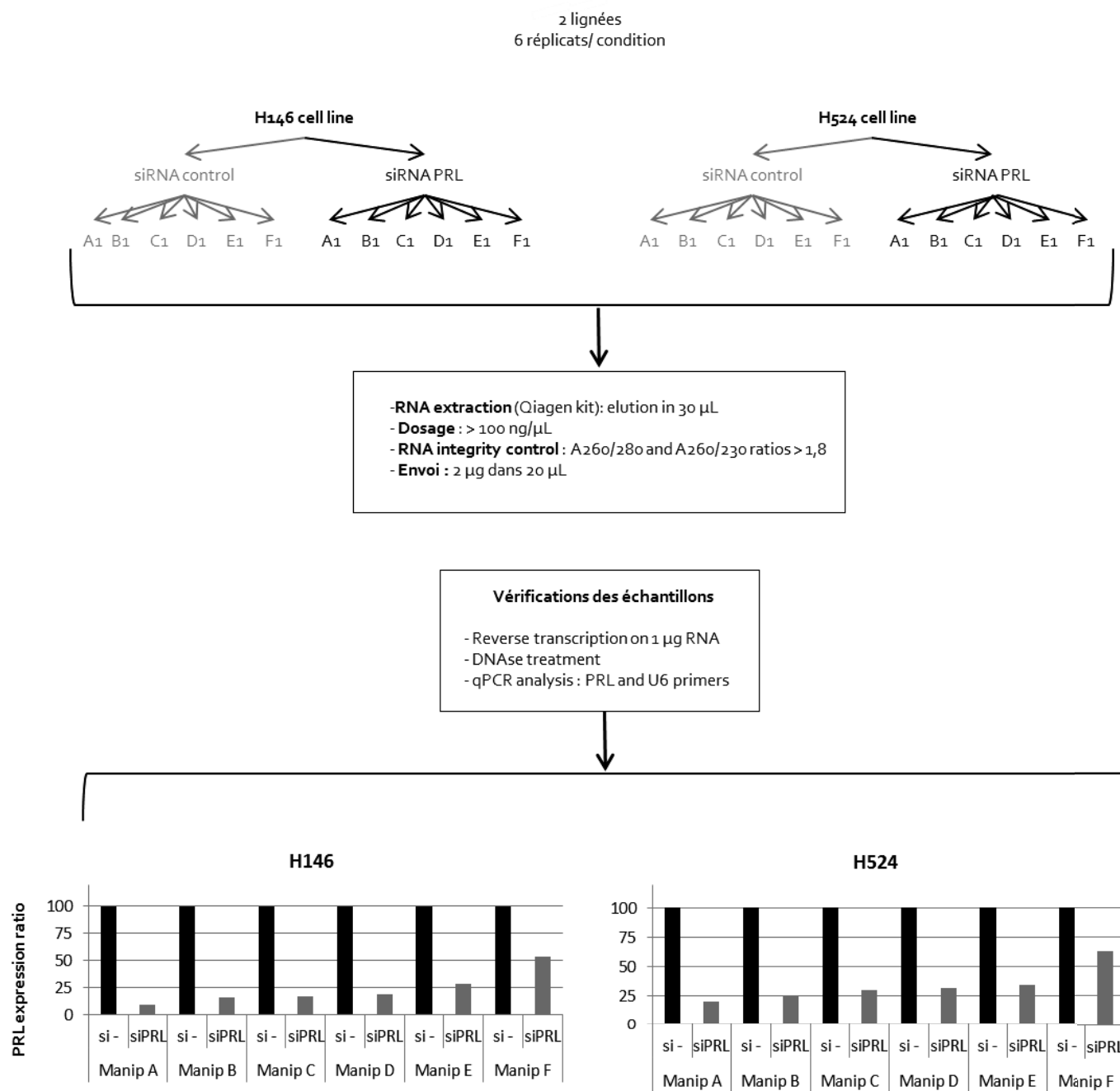


Figure 89 : Echantillons d'ARN des lignées de CBPC H146 et H524 envoyés pour l'analyse transcriptomique

ANNEXES



Contents lists available at ScienceDirect

Biochimica et Biophysica Acta

journal homepage: www.elsevier.com/locate/bbagrm



Review

Induced malignant genome reprogramming in somatic cells by testis-specific factors

Jin Wang^{a,b}, Anouk Emadali^{c,d,e}, Aurore Le Bescont^a, Mary Callanan^{a,f},
Sophie Rousseaux^a, Saadi Khochbin^{a,*}

^a INSERM, U823; Université Joseph Fourier - Grenoble 1; Institut Albert Bonniot, Grenoble, F-38706, France

^b State Key Laboratory of Medical Genomics, Shanghai Institute of Hematology, Department of Hematology, Ruijin Hospital, Shanghai Jiao Tong University School of Medicine, Shanghai 200025, China

^c CEA, IRTSV, Laboratoire Biologie à Grande Echelle, F-38054 Grenoble, France

^d INSERM, U1038, F-38054 Grenoble, France

^e Université Joseph Fourier, Grenoble 1, F-38000, France

^f Onco-Hematology Genetics Laboratory, Plateforme hospitalière de génétique moléculaire des cancers, Grenoble University Hospital Centre, Grenoble, France

ARTICLE INFO

Article history:

Received 4 March 2011

Received in revised form 4 April 2011

Accepted 6 April 2011

Available online 21 April 2011

Keywords:

BRD4-NUT

NUT

BRDT

ATAD2

iPS

ABSTRACT

Germline cell differentiation is controlled by a specific set of genes whose expression is tightly locked into the repressed state in somatic cells. Large-scale epigenome alterations, now evidenced in nearly all cancers, lead to aberrant activation of these normally silenced genes, as attested by the many reports describing the expression of testis-specific factors, known as cancer-testis genes, in various cancer cells. Here, based on the literature, we argue that off-context activity of some of the testis-specific epigenome regulators can reprogram the somatic cell epigenome toward a malignant state by favoring self-renewal and sustaining cell proliferation under stressful conditions, thereby constituting a major oncogenic mechanism.

© 2011 Elsevier B.V. All rights reserved.

1. Introduction

A multistep genome reprogramming process marks the generation and maturation of male germ cells [1–3]. A genome-wide demethylation and remethylation characterizes the first steps of the male genome programming when gene promoters and regulatory elements acquire their male-specific profile of methylation as well as chromatin modifications [1]. It is however the haploid post-meiotic genome that undergoes the most dramatic reprogramming characterized by an almost genome-wide displacement of histones and successive assembly of transition proteins and protamines [4]. Indeed, during this process the universal mode of genome packing in eukaryotes, based on nucleosomes, is drastically altered by transition to a new form of DNA compaction using small basic proteins, called protamines. Although these post-meiotic changes are of key importance for procreation and constitute an essential step in the life cycle, most of the mechanisms controlling the establishment of the male gamete genome remain obscure [5]. Taking into account the published investigations, it is however possible to highlight two essential events in male genome reprogramming which both aim at inducing large-scale destabilization of the “somatic type” genome organization. First, considerable amounts of testis-specific histone variants are synthesized and incorporated in

meiotic and post-meiotic stages. Most of these histones have been shown to induce nucleosome instability. Additionally, just before replacement, a massive histone hyperacetylation also occurs that contributes to chromatin relaxation [4]. Finally, these events as well as the subsequent chromatin disassembly and genome repackaging are hypothesized to be controlled by a variety of testis-specific factors. Although some of them are in the course of identification, many remain to be discovered and functionally characterized [5].

Drastic and genome-wide alterations of the epigenome have now been shown to occur almost without exception in cancer cells [6]. General loss of DNA methylation, as well as of specific histone marks associated with aberrant gain of methylation of some of the CpG islands, appear now as universal hallmarks of oncogenesis [7]. These modifications are expected to profoundly impact gene expression leading to unscheduled gene repression and activation. The oncogenic consequences of aberrant methylation of critical genes such as tumor suppressors have been under intense investigations by many laboratories [8]. In contrast, little attention has been paid to unprogrammed gene activation in cancer cells. Over the last twenty years, investigators looking for tumor specific antigens, have noticed in cancer cells, the abnormal expression of a group of genes which are normally expressed in male germ cells [9]. Some investigations showed that epigenetic mechanisms control their repression in normal somatic cells and concluded hence that their activation reflects the alteration of these mechanisms in somatic cells [10–12].

* Corresponding author. Tel.: +33 4 76 54 95 83; fax: +33 4 76 54 95 95.

E-mail address: khochbin@ujf-grenoble.fr (S. Khochbin).

Despite increasing reports in the literature of testis-specific genes de-repression in many different types of cancers, little is known on the impact of their off-context activity on malignant cell transformation [12].

Since some of these testis-specific genes encode potent epigenetic regulators that play a critical role in the male genome reprogramming [5], we would like to propose that aberrant activation of these factors in the course of somatic cell transformation could target the cell epigenome for abnormal re-programming events that would drive the emergence of fully transformed cancer cell clones.

2. Germ cell epigenetic regulators activated in somatic cancer cells

Although current research is far from providing a comprehensive picture of the epigenetic regulators that mediate the establishment of the germ cell epigenome, some of the determinant factors have been already identified [5] and all of these have been shown to be aberrantly activated in cancer cells [12]. Among these factors, Dnmt3L, a catalytically inactive member of DNA methyltransferases, can be mentioned. Dnmt3L interacts with *de novo* DNA methyltransferases Dnmt3a/3b and by stimulating their catalytic activities, plays an important role in the establishment of DNA methylation on retroposon and that of germ cell-specific genomic loci [13].

Dnmt3L has been shown to be aberrantly activated in a variety of cancers, where due to its ability to enhance the activity of the *de novo* DNA methyltransferases, it could contribute to CpG island methylation and pathological gene silencing [14].

Besides factors involved in DNA methylation, germ cells also express specific histone code readers. Brdt, a double bromodomain-containing protein, is one such factor that is strictly expressed in male germ cells [15]. The protein is capable of specifically acting on hyperacetylated chromatin and mediates a histone acetylation-dependent chromatin compaction [16]. Interestingly, a recent investigation showed that Bdf1, the Brdt ortholog in yeast, is also required for a histone acetylation-dependent genome compaction during the post-meiotic phases of sporulation [17], confirming therefore the particular role for Brdt/Bdf1 in acetylation-dependent chromatin remodeling. Here again, the unscheduled activation of *BRDT* in somatic cells [18] could lead to a large-scale alteration of acetylated chromatin in cancer cells.

Additionally, functional characterizations of several testis-specific factors expressed off-context in somatic cancer cells suggested that they also could function as epigenome regulators in their physiological setting, in the germ cells. One of these factors is NUT (nuclear protein in testis), encoded by a testis-specific gene, aberrantly activated in a subset of very aggressive carcinoma, known as NUT-Midline Carcinoma (NMC) affecting children and adults of different ages [19]. In this cancer, a chromosomal translocation fuses the double bromodomain protein encoded by the ubiquitous gene, *BRD4*, to almost the complete reading frame of *NUT* and hence taking control of its expression. Recent investigations showed that NUT strongly interacts with histone acetyltransferases (HATs) p300 and considerably stimulates its catalytic activity [20]. These findings also suggested that the normal function of NUT in spermatogenic cells could be the enhancement of the HAT activity of p300/CBP, or other HATs, to contribute to the wave of post-meiotic histone hyperacetylation leading to chromatin relaxation and histone replacement. In NMC cells, BRD4-NUT creates a histone super-acetylator, inducing the formation of hyperacetylated chromatin domains due to a combined action of BRD4's bromodomains and the NUT-mediated stimulation of p300 activity, through a feed-forward mechanism completely affecting the epigenetic landscape of the affected cells [20].

Another factor predominantly expressed in testis and capable of acting on chromatin is ATAD2. The protein harbors an AAA ATPase domain responsible for its multimerization and chromatin binding, as well as a bromodomain. This bromodomain, at least in a testis-specific splice variant, has been shown to specifically bind histone H4 bearing

acetyl-K5. ATAD2 appears to control chromatin dynamics and could also play an important role in male genome reprogramming in post-meiotic stages, when it is abundantly expressed [21]. Surprisingly, ATAD2 is activated in a great number of unrelated somatic cancers and in several cases its over-expression has been shown to correlate with a rapidly fatal outcome [21–23].

Finally, BORIS, a testis-specific paralog of CTCF, has been recently shown to be a critical regulator of other testis-specific genes required for normal spermatogenesis [24]. The illegitimate activation of BORIS affects somatic cell gene expression programs and in some cases could turn on several normally silenced testis-specific genes [25–29].

There are very probably many other important germ cell-specific epigenetic regulators activated in cancer cells but in the absence of more functional insights into the nature of factors involved in male genome reprogramming, they are at the present not identifiable.

3. Cancer-testis factors as mediators of critical transcriptional regulatory circuits in somatic cells

The off-context activity of the above-mentioned factors has the potential to profoundly and globally affect the cell epigenome but there are also several recent investigations reporting the ability of cancer-testis factors to alter the activity of critical cellular transcription factors and hence a specific downstream regulatory pathway. The majority of data on this aspect concerns proteins encoded by the *MAGE* family of cancer-testis genes as regulators of p53 activity. Early works showed that several members of the *MAGE* family can suppress p53-dependent apoptosis in melanoma cell lines [30]. Detailed investigations unraveled the capacity of a specific member, *MAGE-A2*, to recruit HDAC3 and to mediate p53 deacetylation and inactivation after a genotoxic treatment [31]. Additionally, a newly discovered function of *MAGE* family members is the stimulation of RING E3 ubiquitin ligases, which could be responsible for proteasomal degradation of a variety of critical cellular factors. With this respect, the *MAGE-C2-TRIM28* complex has been shown to induce ubiquitination and degradation of p53 [32].

In addition to modulating p53 activity, *MAGE* family members can also affect several unrelated transcriptional regulatory circuits such as those controlled by the androgen receptor (AR). Indeed, *MAGE-11*, another member of the *MAGE* family, has the ability to stimulate the transcriptional activity of this receptor [33]. A master transcription factor also controlled by *MAGE* members is Hif1- α . *MAGE-11*, as well as a closely related member, *MAGE-9*, interact with the major Hif1- α hydroxylating enzyme, PHD2, and reduce its catalytic activity. Prolyl-hydroxylases mediate site-specific hydroxylation of Hif1- α and induce its degradation by proteasome. The aberrant expression of *MAGE-11/9* in tumors therefore leads to the stabilization of Hif1- α and the subsequent stimulation of angiogenesis and glycolytic switch [34]. Finally, *MAGE-A1* through the recruitment of HDAC1 can turn a transcriptional activator, SKIP, into a repressor and hence modify the state of cancer cell gene expression [35].

The p53 gene appears however as a major target of the off-context activity of the C/T factors. Indeed, another member of C/T factors, *CAGE* has been recently shown to use HDAC2 to repress p53 gene expression in melanoma cell lines and thus confers resistance to drug-induced apoptosis [36]. The off-context activity of NUT, in fusion with BRD4 also affects p53 functions through the sequestration of CBP/p300 into BRD4-NUT foci, which hinders the activation of p53 in response to genotoxic treatments [20].

Like *MAGE* members, both NUT and *CAGE* also control critical transcriptional circuits. A recent work showed that a matrix metalloproteinase, MMP-2 could be induced through the action of *CAGE* on the NF- κ B and AP1 transcriptional pathways [37]. NUT in fusion with BRD4 titers out CBP/p300 and has the potential to affect the expression of numerous genes controlled by these HATs [20].

Cancer-testis factors could also turn on a cell proliferation program. Indeed, the C/T factor CAGE, has been reported to use AP-1 and E2F-1 transcription factors to activate D1 and E-type cyclins, thus enhancing G1 to S phase progression [38]. Interestingly, ATAD2 also turns on a proliferation program in cancer cells upon its aberrant activation. ATAD2 activity as a master regulator of cell proliferation is two fold: it cooperates with critical factors such as Myc [22] and E2F-1 [39] and it specifically or globally affects gene expression by modifying the expression of MLL and EZH2 histone methyltransferases.

Finally, the aberrant activation of testis-specific small RNA systems in cancer cells also represents an important mechanism for global or specific alterations of gene expression in transformed cells. For instance, *PIWIL2*, a member of PIWI/AGO gene family, specifically expressed in testis, has been shown to be ectopically expressed in a variety of cancers [40–42]. Interestingly, *PiwiL2* behaves as a potent oncogene and its ectopic expression alone is sufficient to induce the transformation of fibroblasts using a mechanism involving the anti-apoptotic factor Bcl-X(L) and Stat3 transcription factor [41,43].

3.1. Ectopic expression of testis-specific genes derives a soma-to-germline transformation

As discussed in this review, several key transcriptional regulators such as ATAD2, CAGE and PIWIL2 could activate a cell proliferation-related gene expression program, but taking into account the properties of some the testis-specific epigenetic factors activated in cancer cells, one can hypothesize that their action could also turn on a specific developmental program. This hypothesis has very recently received important support from a work describing the ectopic expression of germline genes in malignant brain tumors generated in drosophila [44]. Indeed, a detailed analysis of cancer-related gene expression in drosophila brain tumors induced after the inactivation of tumor suppressor gene, *l(3)mbt*, showed that about 25% of the activated genes belong to the germ cell transcript repertoire. Several of these genes encode proteins involved in piRNA pathways and probably account for the ectopic expression of small RNAs, which are

normally present in ovaries, in these brain tumors. Interestingly, *vasa*, a component of a germ cell RNA-processing organelle, *nuage*, is among the ectopically expressed genes. This is highly relevant to the activation of germ cell gene expression, since it has recently been shown that the ectopic expression of the chicken *vasa* homolog, *Cvh*, mediates the reprogramming of chicken embryonic stem cells to germ cell fate [45]. This action of *Cvh* was strongly stimulated when a testis-specific microRNA, miR-34c, was co-expressed with *Cvh*. Additionally, miR-34c seems to be sufficient to turn on a germline gene expression program when expressed alone in HeLa cells [46].

Here again, as in the case of the C/T factors discussed above, several of these germ cell genes ectopically expressed in the drosophila brain tumors, have been shown to be responsible for the tumor growth, indicating that in addition to the germline genes they can also turn on a proliferation program.

Although it appears beyond doubt that the ectopic expression of C/T factors contributes to the cancer-specific pattern of gene expression, it has not been clearly established how and why the germline gene expression program becomes activated in cancer cells. In the case of drosophila [44] as well as *C. elegans* [47,48], the retinoblastoma pathway seems to be important in the repression of germline genes in somatic cells. In vertebrates, the epigenetic landscape, mainly DNA methylation and histone H3 K9 methylation appear important in the repression of these genes [6,12,49]. Germline-specific gene promoter methylation could however be dependent on the activity of a particular transcription factor, E2F6, an Rb-independent transcriptional repressor, which has been shown to be critical in the restriction of testis-specific genes in somatic cells [50] through recognition of a motif common to germline genes and induction of DNA methylation through the action of Dnmt3b [51]. Post-translational SUMO modification of a particular transcription factor, Sp3, has also been shown to be responsible for the repression of spermatocyte-specific genes in somatic cells, which somehow also controls the methylation of these genes [52].

These data suggest that in addition to global epigenetic alterations, the deregulated function of a group of specific factors could lead to the activation of a significant number of germline genes in somatic cells.

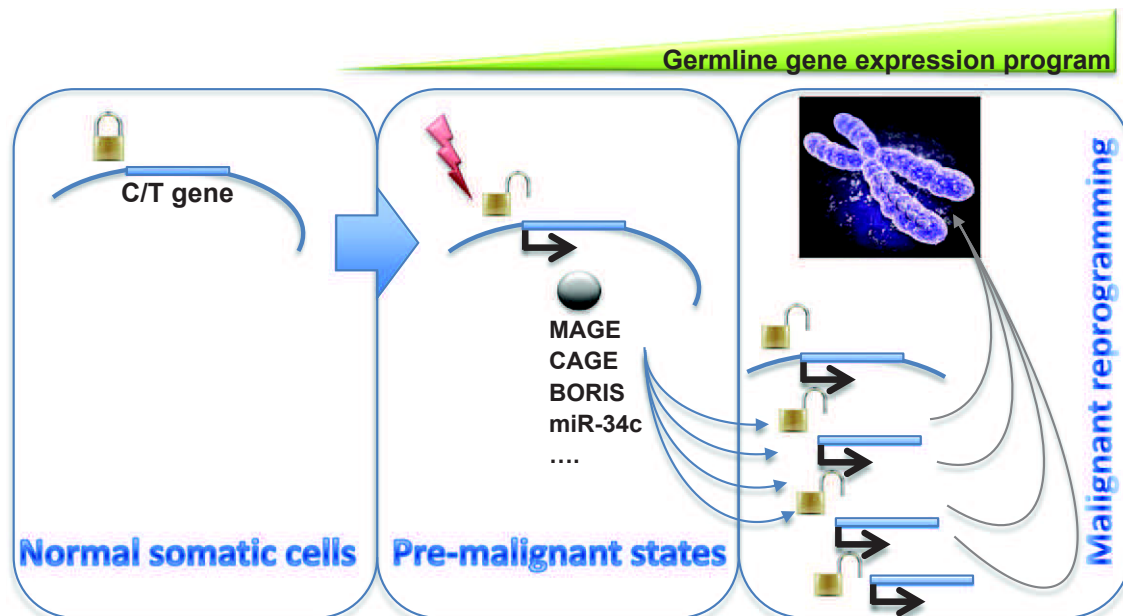


Fig. 1. Somatic cell genome reprogramming to the malignant state by aberrant activation of testis-specific factors. Under transformation-prone conditions, very limited or subtle alterations of gene silencing due to abnormal epigenetic signaling or deregulated activity of specific factors could activate one or several cancer-testis (C/T) gene members whose expression is normally tightly locked into a silenced state in somatic cells. Off-context activity of these factors on critical regulatory circuits would activate anti-apoptotic and proliferative gene expression programs. De-repression of testis-specific epigenome regulators within the C/Ts gene family could further extend epigenetic deregulations in these cells and initiate an amplification process leading to activation of more C/Ts and accentuation of epigenetic alterations leading to complete reprogramming of the affected cell epigenome, toward the malignant state. For an overview of all the known C/T genes and insight into their function, please refer to: <http://www.cta.lncc.br/index.php>.

3.2. Concluding remarks

A direct consequence of large-scale epigenetic alterations now reported in almost all cancers, is the induction of an “identity crisis” in the affected cells due to inappropriate activation of tissue-restricted genes. Analysis of the literature indicates however that there is a tendency toward the activation of the germ cell genetic program in malignant cells [44]. Based on this observation, it can be suggested that germline-specific genes confer a selective advantage in the process of oncogenic transformation. The fact that a subset of testis-specific factors constitutes very potent epigenome regulators, could at least in part explain why the ectopic expression of testis-specific genes could confer a selective advantage.

Indeed, the expression of factors such as testis-specific histone variants, which almost systematically induce nucleosome instability [4], could trigger secondary alterations, globally affecting genome organization and the state of epigenome. This is also true for numerous other factors controlling male germ cell genome methylation and chromatin remodeling. An amplification loop mechanism could be hypothesized postulating that limited initial alterations in gene expression leading to the derepression of potent testis-specific epigenome regulators could initiate a cascade of aberrant gene activations due to successive waves of gene activation/repression and epigenome alteration cycles (Fig. 1). The germline program could then become a preferential target for activation, since most of the corresponding genes present a very distinct epigenetic status in somatic cells [53], and would thus constitute a pool of genes that would be particularly sensitive to global epigenetic alterations.

Additionally, some of the testis-specific factors could act as master regulators controlling the expression of other germline genes. An example is the ectopic activation of *miR-34c* and *vasa*, which in turn could activate many other spermatogenic genes [45,46]. The primary deregulation of such genes could therefore also account for the turning on of a spermatogenic program, which in the context of somatic cells, would confer considerable growth advantages particularly under highly unfavorable or stressful conditions.

Acknowledgments

This review corresponds to a research project supported by INCa-DHOS in J. Garin, MC and SK laboratories and by ARC and ANR in SK laboratory. AE is a recipient of a post-doctoral fellowship by INCa-DHOS (JG laboratory). JW's work is supported by a grant from the Science and Technology Commission of Shanghai Municipality (09540700800).

References

- [1] S. Rousseaux, C. Caron, J. Govin, C. Lestrat, A.-K. Faure, S. Khochbin, Establishment of male-specific epigenetic information, *Gene* 345 (2005) 139–153.
- [2] F. Boussouar, S. Rousseaux, S. Khochbin, A new insight into male genome reprogramming by histone variants and histone code, *Cell Cycle* 7 (2008) 3499–3502.
- [3] J. Govin, C. Caron, C. Lestrat, S. Rousseaux, S. Khochbin, The role of histones in chromatin remodelling during mammalian spermiogenesis, *Eur. J. Biochem.* 271 (2004) 3459–3469.
- [4] J. Gaucher, N. Reynold, E. Montellier, F. Boussouar, S. Rousseaux, S. Khochbin, From meiosis to postmeiotic events: the secrets of histone disappearance, *FEBS J.* 277 (2010) 599–604.
- [5] S. Rousseaux, F. Boussouar, J. Gaucher, N. Reynold, E. Montellier, S. Curtet, A.-L. Vitte, S. Khochbin, Molecular models for post-meiotic male genome reprogramming, *Syst. Biol. Reprod. Med.* 57 (2011) 50–53.
- [6] M. Berdasco, M. Esteller, Aberrant epigenetic landscape in cancer: how cellular identity goes awry, *Dev. Cell* 19 (2010) 698–711.
- [7] M. Esteller, Cancer epigenomics: DNA methylomes and histone-modification maps, *Nat. Rev. Genet.* 8 (2007) 286–298.
- [8] N. Sincic, Z. Herceg, DNA methylation and cancer: ghosts and angels above the genes, *Curr. Opin. Oncol.* 23 (2011) 69–76.
- [9] A.J.G. Simpson, O.L. Caballero, A. Jungbluth, Y.-T. Chen, L.J. Old, Cancer/testis antigens, gametogenesis and cancer, *Nat. Rev. Cancer* 5 (2005) 615–625.
- [10] A.R. Karpf, D.A. Jones, Reactivating the expression of methylation silenced genes in human cancer, *Oncogene* 21 (2002) 5496–5503.
- [11] A.R. Karpf, Epigenomic reactivation screening to identify genes silenced by DNA hypermethylation in human cancer, *Curr. Opin. Mol. Ther.* 9 (2007) 231–241.
- [12] S. Rousseaux, S. Khochbin, New hypotheses for large-scale epigenome alterations in somatic cancer cells: a role for male germ-cell-specific regulators, *Epigenomics* 1 (2009) 153–161.
- [13] R.J. Klose, A.P. Bird, Genomic DNA methylation: the mark and its mediators, *Trends Biochem. Sci.* 31 (2006) 89–97.
- [14] G. Gokul, G. Ramakrishna, S. Khosla, Reprogramming of HeLa cells upon DNMT3L overexpression mimics carcinogenesis, *Epigenetics* 4 (2009) 322–329.
- [15] C. Pivot-Pajot, C. Caron, J. Govin, A. Vion, S. Rousseaux, S. Khochbin, Acetylation-dependent chromatin reorganization by BRDT, a testis-specific bromodomain-containing protein, *Mol. Cell. Biol.* 23 (2003) 5354–5365.
- [16] J. Morinière, S. Rousseaux, U. Steuerwald, M. Soler-López, S. Curtet, A.-L. Vitte, J. Govin, J. Gaucher, K. Sadoul, D.J. Hart, J. Krijgsvel, S. Khochbin, C.W. Müller, C. Petosa, Cooperative binding of two acetylation marks on a histone tail by a single bromodomain, *Nature* 461 (2009) 664–668.
- [17] J. Govin, J. Schug, T. Krishnamoorthy, J. Dorsey, S. Khochbin, S.L. Berger, Genome-wide mapping of histone H4 serine-1 phosphorylation during sporulation in *Saccharomyces cerevisiae*, *Nucleic Acids Res.* 38 (2010) 4599–4606.
- [18] M.J. Scanlan, N.K. Altorki, A.O. Gure, B. Williamson, A. Jungbluth, Y.T. Chen, L.J. Old, Expression of cancer-testis antigens in lung cancer: definition of bromodomain testis-specific gene (BRDT) as a new CT gene, *Cancer Lett.* 150 (2000) 155–164.
- [19] C.A. French, C.L. Ramirez, J. Kolmakova, T.T. Hickman, M.J. Cameron, M.E. Müller, J. L. Kutok, J.A. Toretsky, A.K. Tadavarthy, U.R. Kees, J.A. Fletcher, J.C. Aster, BRD-NUT oncoproteins: a family of closely related nuclear proteins that block epithelial differentiation and maintain the growth of carcinoma cells, *Oncogene* 27 (2008) 2237–2242.
- [20] N. Reynold, B.E. Schwartz, M. Delvecchio, K. Sadoul, D. Meyers, C. Mukherjee, C. Caron, H. Kimura, S. Rousseaux, P.A. Cole, D. Panne, C.A. French, S. Khochbin, Oncogenesis by sequestration of CBP/p300 in transcriptionally inactive hyperacetylated chromatin domains, *EMBO J.* 29 (2010) 2943–2952.
- [21] C. Caron, C. Lestrat, S. Marsal, E. Escoffier, S. Curtet, V. Virolle, P. Barbry, A. Debernardi, C. Brambilla, E. Brambilla, S. Rousseaux, S. Khochbin, Functional characterization of ATAD2 as a new cancer/testis factor and a predictor of poor prognosis in breast and lung cancers, *Oncogene* 29 (2010) 5171–5181.
- [22] M. Ciró, E. Prosperini, M. Quarto, U. Grazini, J. Walfridsson, F. McBlane, P. Nucifero, G. Pacchiana, M. Capra, J. Christensen, K. Helin, ATAD2 is a novel cofactor for MYC, overexpressed and amplified in aggressive tumors, *Cancer Res.* 69 (2009) 8491–8498.
- [23] E.V. Kalashnikova, A.S. Revenko, A.T. Gemo, N.P. Andrews, C.G. Tepper, J.X. Zou, R. D. Cardiff, A.D. Borowsky, H.-W. Chen, ANCCA/ATAD2 overexpression identifies breast cancer patients with poor prognosis, acting to drive proliferation and survival of triple-negative cells through control of B-Myb and E2H2, *Cancer Res.* 70 (2010) 9402–9412.
- [24] T. Suzuki, N. Kosaka-Suzuki, S. Pack, D.M. Shin, J. Yoon, Z. Abdullaev, E. Pugacheva, H.C. Morse III, D. Loukinov, V. Lobanenko, Expression of a testis-specific form of Gal3st1 (CST), a gene essential for spermatogenesis, is regulated by the CTCF paralogous gene BORIS, *Mol. Cell. Biol.* 30 (2010) 2473–2484.
- [25] E.M. Klenova, H.C. Morse, R. Ohlsson, V.V. Lobanenko, The novel BORIS + CTCF gene family is uniquely involved in the epigenetics of normal biology and cancer, *Semin. Cancer Biol.* 12 (2002) 399–414.
- [26] L. Sun, L. Huang, P. Nguyen, K.S. Bisht, G. Bar-Sela, A.S. Ho, C.M. Bradbury, W. Yu, H. Cui, S. Lee, J.B. Trepel, A.P. Feinberg, D. Gius, DNA methyltransferase 1 and 3B activate BAG-1 expression via recruitment of CTCF/BORIS and modulation of promoter histone methylation, *Cancer Res.* 68 (2008) 2726–2735.
- [27] S. Vatolin, Z. Abdullaev, S.D. Pack, P.T. Flanagan, M. Custer, D.I. Loukinov, E. Pugacheva, J.A. Hong, H. Morse III, D.S. Schrupp, J.I. Risinger, J.C. Barrett, V.V. Lobanenko, Conditional expression of the CTCF-paralogous transcriptional factor BORIS in normal cells results in demethylation and derepression of MAGE-A1 and reactivation of other cancer-testis genes, *Cancer Res.* 65 (2005) 7751–7762.
- [28] J.A. Hong, Y. Kang, Z. Abdullaev, P.T. Flanagan, S.D. Pack, M.R. Fischette, M.T. Adnani, D.I. Loukinov, S. Vatolin, J.I. Risinger, M. Custer, G.A. Chen, M. Zhao, D.M. Nguyen, J.C. Barrett, V.V. Lobanenko, D.S. Schrupp, Reciprocal binding of CTCF and BORIS to the NY-ESO-1 promoter coincides with derepression of this cancer-testis gene in lung cancer cells, *Cancer Res.* 65 (2005) 7763–7774.
- [29] Y. Kang, J.A. Hong, G.A. Chen, D.M. Nguyen, D.S. Schrupp, Dynamic transcriptional regulatory complexes including BORIS, CTCF and Sp1 modulate NY-ESO-1 expression in lung cancer cells, *Oncogene* 26 (2007) 4394–4403.
- [30] B. Yang, S. O'Herrin, J. Wu, S. Reagan-Shaw, Y. Ma, M. Nihal, B.J. Longley, Select cancer testis antigens of the MAGE-A, -B, and -C families are expressed in mast cell lines and promote cell viability in vitro and in vivo, *J. Invest. Dermatol.* 127 (2007) 267–275.
- [31] M. Monte, M. Simonatto, L.Y. Pêche, D.R. Bublik, S. Gobessi, M.A. Pierotti, M. Rodolfo, C. Schneider, MAGE-A tumor antigens target p53 transactivation function through histone deacetylase recruitment and confer resistance to chemotherapeutic agents, *Proc. Natl. Acad. Sci. U. S. A.* 103 (2006) 11160–11165.
- [32] J.M. Doyle, J. Gao, J. Wang, M. Yang, P.R. Potts, MAGE-RING protein complexes comprise a family of E3 ubiquitin ligases, *Mol. Cell* 39 (2010) 963–974.
- [33] S. Bai, E.M. Wilson, Epidermal-growth-factor-dependent phosphorylation and ubiquitinylation of MAGE-11 regulates its interaction with the androgen receptor, *Mol. Cell. Biol.* 28 (2008) 1947–1963.
- [34] O. Aprelikova, S. Pandolfi, S. Tackett, M. Ferreira, K. Salnikow, Y. Ward, J.I. Risinger, J.C. Barrett, J. Niederhuber, Melanoma antigen-11 inhibits the hypoxia-inducible factor prolyl hydroxylase 2 and activates hypoxic response, *Cancer Res.* 69 (2009) 616–624.

- [35] S. Laduron, R. Deplus, S. Zhou, O. Kholmanskikh, D. Godelaine, C. De Smet, S.D. Hayward, F. Fuks, T. Boon, E. De Plaen, MAGE-A1 interacts with adaptor SKIP and the deacetylase HDAC1 to repress transcription, *Nucleic Acids Res.* 32 (2004) 4340–4350.
- [36] Y. Kim, H. Park, D. Park, Y.-S. Lee, J. Choe, J.-H. Hahn, H. Lee, Y.-M. Kim, D. Jeoung, Cancer/testis antigen CAGE exerts negative regulation on p53 expression through HDAC2 and confers resistance to anti-cancer drugs, *J. Biol. Chem.* 285 (2010) 25957–25968.
- [37] Y. Kim, D. Jeoung, The cancer/testis antigen CAGE induces MMP-2 through the activation of NF-kappaB and AP-1, *BMB Rep.* 42 (2009) 758–763.
- [38] E. Por, H.-J. Byun, E.-J. Lee, J.-H. Lim, S.-Y. Jung, I. Park, Y.-M. Kim, D.-I. Jeoung, H. Lee, The cancer/testis antigen CAGE with oncogenic potential stimulates cell proliferation by up-regulating cyclins D1 and E in an AP-1- and E2F-dependent manner, *J. Biol. Chem.* 285 (2010) 14475–14485.
- [39] A.S. Revenko, E.V. Kalashnikova, A.T. Gemo, J.X. Zou, H.-W. Chen, Chromatin loading of E2F-MLL complex by cancer-associated coregulator ANCCA via reading a specific histone mark, *Mol. Cell. Biol.* 30 (2010) 5260–5272.
- [40] J.J. Liu, R. Shen, L. Chen, Y. Ye, G. He, K. Hua, D. Jarjoura, T. Nakano, G.K. Ramesh, C.L. Shapiro, S.H. Barsky, J.-X. Gao, Piwil2 is expressed in various stages of breast cancers and has the potential to be used as a novel biomarker, *Int. J. Clin. Exp. Pathol.* 3 (2010) 328–337.
- [41] J.H. Lee, D. Schütte, G. Wulf, L. Füzesi, H.-J. Radzun, S. Schweyer, W. Engel, K. Nayermia, Stem-cell protein Piwil2 is widely expressed in tumors and inhibits apoptosis through activation of Stat3/Bcl-XL pathway, *Hum. Mol. Genet.* 15 (2006) 201–211.
- [42] G. He, L. Chen, Y. Ye, Y. Xiao, K. Hua, D. Jarjoura, T. Nakano, S.H. Barsky, R. Shen, J.-X. Gao, Piwil2 expressed in various stages of cervical neoplasia is a potential complementary marker for p16, *Am. J. Transl. Res.* 2 (2010) 156–169.
- [43] Y. Ye, D.T. Yin, L. Chen, Q. Zhou, R. Shen, G. He, Q. Yan, Z. Tong, A.C. Issekutz, C.L. Shapiro, S.H. Barsky, H. Lin, J.J. Li, J.X. Gao, Identification of Piwil2-like (PL2L) proteins that promote tumorigenesis, *PLoS One* 5 (2010) e13406.
- [44] A. Janic, L. Mendizabal, S. Llamazares, D. Rossell, C. Gonzalez, Ectopic expression of germline genes drives malignant brain tumor growth in *Drosophila*, *Science* 330 (2010) 1824–1827.
- [45] F. Lavial, H. Acloque, E. Bachelard, M.A. Nieto, J. Samarut, B. Pain, Ectopic expression of Cvh (Chicken Vasa homologue) mediates the reprogramming of chicken embryonic stem cells to a germ cell fate, *Dev. Biol.* 330 (2009) 73–82.
- [46] F. Bouhallier, N. Alloli, F. Lavial, F. Chalmel, M.-H. Perrard, P. Durand, J. Samarut, B. Pain, J.-P. Rouault, Role of miR-34c microRNA in the late steps of spermatogenesis, *RNA* 16 (2010) 720–731.
- [47] D. Wang, S. Kennedy, D. Conte Jr., J.K. Kim, H.W. Gabel, R.S. Kamath, C.C. Mello, G. Ruvkun, Somatic misexpression of germline P granules and enhanced RNA interference in retinoblastoma pathway mutants, *Nature* 436 (2005) 593–597.
- [48] Y. Unhavaithaya, T.H. Shin, N. Miliaras, J. Lee, T. Oyama, C.C. Mello, MEP-1 and a homolog of the NURD complex component Mi-2 act together to maintain germline-soma distinctions in *C. elegans*, *Cell* 111 (2002) 991–1002.
- [49] A.R. Karpf, A potential role for epigenetic modulatory drugs in the enhancement of cancer/germ-line antigen vaccine efficacy, *Epigenetics* 1 (2006) 116–120.
- [50] M. Pohlers, M. Truss, U. Frede, A. Scholz, M. Strehle, R.-J. Kuban, B. Hoffmann, M. Morkel, C. Birchmeier, C. Hagemeier, A role for E2F6 in the restriction of male-germ-cell-specific gene expression, *Curr. Biol.* 15 (2005) 1051–1057.
- [51] G. Velasco, F. Hube, J. Rollin, D. Neuillet, C. Philippe, H. Bouzinba-Segard, A. Galvani, E. Viegas-Pequignot, C. Francastel, Dnmt3b recruitment through E2F6 transcriptional repressor mediates germ-line gene silencing in murine somatic tissues, *Proc. Natl. Acad. Sci. U. S. A.* 107 (2010) 9281–9286.
- [52] B. Stielow, I. Kruger, R. Diezko, F. Finkernagel, N. Gillemans, J. Kong-a-San, S. Philipsen, G. Suske, Epigenetic silencing of spermatocyte-specific and neuronal genes by SUMO modification of the transcription factor Sp3, *PLoS Genet.* 6 (2010) e1001203.
- [53] M. Weber, I. Hellmann, M.B. Stadler, L. Ramos, S. Paabo, M. Rebhan, D. Schubeler, Distribution, silencing potential and evolutionary impact of promoter DNA methylation in the human genome, *Nat. Genet.* 39 (2007) 457–466.

LUNG CANCER

Ectopic Activation of Germline and Placental Genes Identifies Aggressive Metastasis-Prone Lung Cancers

Sophie Rousseaux,^{1*} Alexandra Debernardi,¹ Baptiste Jacquiau,¹ Anne-Laure Vitte,¹ Aurélien Vesin,¹ Hélène Nagy-Mignotte,² Denis Moro-Sibilot,² Pierre-Yves Brichon,² Sylvie Lantuejoul,² Pierre Hainaut,³ Julien Laffaire,⁴ Aurélien de Reyniès,⁴ David G. Beer,⁵ Jean-François Timsit,^{1,2} Christian Brambilla,^{1,2} Elisabeth Brambilla,^{1,2} Saadi Khochbin^{1*}

Activation of normally silent tissue-specific genes and the resulting cell “identity crisis” are the unexplored consequences of malignant epigenetic reprogramming. We designed a strategy for investigating this reprogramming, which consisted of identifying a large number of tissue-restricted genes that are epigenetically silenced in normal somatic cells and then detecting their expression in cancer. This approach led to the demonstration that large-scale “off-context” gene activations systematically occur in a variety of cancer types. In our series of 293 lung tumors, we identified an ectopic gene expression signature associated with a subset of highly aggressive tumors, which predicted poor prognosis independently of the TNM (tumor size, node positivity, and metastasis) stage or histological subtype. The ability to isolate these tumors allowed us to reveal their common molecular features characterized by the acquisition of embryonic stem cell/germ cell gene expression profiles and the down-regulation of immune response genes. The methodical recognition of ectopic gene activations in cancer cells could serve as a basis for gene signature-guided tumor stratification, as well as for the discovery of oncogenic mechanisms, and expand the understanding of the biology of very aggressive tumors.

INTRODUCTION

Cell differentiation is associated with the establishment of specific patterns of cell type- and tissue-specific gene expression, which largely rely on the cell’s “epigenetic landscape,” mainly shaped by chemical modifications of the genome and the associated histones. In differentiated cells, these epigenetic mechanisms not only help activate and maintain specific gene expression patterns but also control a genome-wide repression of tissue-specific genes (1, 2).

Recent investigations have demonstrated that a global deregulation of epigenetic signaling is an early and recurrent event that occurs during oncogenic cell transformation. Aberrant gene activity is a direct consequence of these anomalies. DNA methylation-associated repression of tumor suppressor genes in cancer cells is well documented and now recognized as an important oncogenic event (3, 4). A less studied consequence of epigenetic deregulations in transformed cells is the ectopic activation of various cell- and tissue-specific genes (5, 6). The aberrant activation of genes in cancer represents a promising source of cancer biomarkers, as exemplified by the discovery of an almost universal cancer marker, the follicle-stimulating hormone (FSH) receptor, abnormally expressed in many types of cancers (7). However, these unprogrammed gene activation events have to date been only sporadically discovered.

Among these genes, male germ cell-specific genes are of particular interest. Indeed, these genes, normally exclusively expressed in testis, have been found to be sporadically derepressed in several somatic cancers and have hence been called cancer/testis (C/T) genes (8),

and proposed as candidate cancer markers with additional potential for exploitation as novel therapeutic targets (C/T gene products are highly immunogenic). However, the sporadic nature of their ectopic expression hinders their use as reliable cancer indicators when detected individually.

The identification of genes with a highly specific pattern of expression, whose silencing was associated with a specific epigenetic signature in normal adult somatic tissues, served as a basis for a genome-wide inventory of all tissue-specific genes with aberrant activations (“off context”) in cancer cells. We show here that this approach provides us with access to a large number of candidate cancer biomarkers and sheds new insight upon the biology of cancer with applications in translational medicine.

RESULTS

Most tissue-restricted genes are epigenetically marked germline genes

To evaluate the extent of ectopic gene activations in cancer cells, it was necessary to first identify the human genes whose physiological pattern of expression is normally restricted to one tissue.

To this end, we combined two strategies, exploiting human ESTs (expressed sequence tags) and transcriptomic data sets. A list of genes whose ESTs were specifically found in a given tissue was established. In parallel, human tissue transcriptomic data (normal tissue samples listed in table S1) were analyzed to isolate a second list of genes with a clearly predominant expression in one tissue (meaning, that gene expression in this tissue was more than 3 SDs above the mean of the values of expression in all tissues). The two gene lists were then compared, and genes that were common to both lists were selected and defined as tissue-specific (fig. S1A, step 1).

Both EST and transcriptomic approaches revealed that testis and germline tissues show the highest number of genes satisfying our

¹INSERM, U823; Université Joseph Fourier, Grenoble 1; Institut Albert Bonniot, Grenoble F-38700, France. ²Grenoble University Hospital (CHU), Grenoble F-38700, France. ³International Prevention Research Institute, Lyon F-69006, France. ⁴Ligue Nationale Contre le Cancer, Cartes d'Identité des Tumeurs Program, Paris F-75013, France. ⁵Thoracic Surgery, University of Michigan, 6304 Cancer Center, Ann Arbor, MI 48109, USA.

*Corresponding author. E-mail: sophie.rousseau@ujf-grenoble.fr (S.R.); khochbin@ujf-grenoble.fr (S.K.)

selection criteria for strict tissue specificity (fig. S1, B and C, respectively), which also explains why most of the already known off-context expressed genes are testis genes (8). Only germline- and placenta-specific genes were selected as potential biomarkers, because they are the only ones that are never expressed in the healthy somatic tissues of an adult organism and remain unknown to the immune system. A total of 506 tissue-restricted genes, including 439 germline- and 67 placenta-restricted genes, were identified by the present approach (hereinafter named TS and PS genes, respectively; listed in table S2).

To test our hypothesis that the single tissue-restricted expression of testis-specific/placenta-specific (TS/PS) genes and their consistently repressed state in adult somatic tissues could be linked to a particular epigenetic status in normal cells, we interrogated several genome-wide epigenetic mapping data sets to characterize the status of these genes (fig. S1A, step 2). Using methylated DNA immunoprecipitation and chromatin immunoprecipitation data from published genome-wide studies (9–11), we extracted epigenomic data corresponding to our list of TS/PS genes. This analysis demonstrated that most of the TS/PS genes show particular promoter sequence features and are associated with a specific epigenetic status in somatic cells, where their expression is tightly locked. In particular, most TS/PS genes are associated with either CpG-poor promoters (46%) or CpG-rich, but hypermethylated, promoters (35%), whereas when all human genes are considered, the proportions of CpG-poor and hypermethylated CpG-rich promoters drop to 22 and 6%, respectively (fig. S1D). We also screened our TS/PS genes for specific transcriptional/epigenetic features in different somatic tissues from various genome-wide studies (10, 11). We found that TS/PS genes are nearly all devoid of active histone marks, including acetylation, H3K4 methylation, and binding of RNA polymerase II, but are relatively enriched in repressive histone modification H3K9 and H3K27 methylation (fig. S1E). The epigenetic status of these TS/PS genes is therefore consistent with our expression data and confirms that we have isolated a specific fraction of human genes. The other observation from these analyses is the homogeneous and consistent nature of the TS/PS gene epigenetic signature, which seems independent of the tissue type, because it was found both in embryonic stem cells and in multiple types of differentiated adult cells, and was reproducible between multiple independent studies (fig. S1E).

Large-scale activations of TS/PS genes occur in all cancer cells

We systematically investigated the expression status of TS/PS genes in human cancers. The analysis of a large set of transcriptomic data from 1776 solid tumor samples derived from 14 different cancer types (described in table S3) showed that hundreds of these genes are aberrantly activated across all cancer types tested (Fig. 1A and table S4), therefore providing access to an invaluable source of cancer biomarkers.

Before going further with the analyses of ectopic expressions of the TS/PS genes in our list, and to verify the possibility of using them as cancer biomarkers, we proceeded to a series of “wet-bench” validations. We designed a custom-dedicated microarray with probes corresponding to our TS/PS gene list. RNAs from 18 nontumor tissues, including testis and placenta as well as somatic tissues, were probed accordingly. As expected, all the tested genes were only expressed in testis (Fig. 1B, “T”) and placenta (Fig. 1B, “P”), and no measurable signal was observed when analyzing RNA from normal tissues (Fig. 1B, “Ctrl soma”). Additionally, we tested RNAs from 21 cancer samples (Fig. 1B, “Cancer”) corresponding to eight different cancer types as well as three cancer cell lines (Fig. 1B, “CCL”) (all samples in our

dedicated array are listed in table S5). The data confirmed that TS/PS genes were ectopically activated in all tested cancer samples (Fig. 1B).

An additional validation was carried out using a quantitative reverse transcription polymerase chain reaction (qRT-PCR) assay testing the activation of 13 of our TS/PS genes, arbitrarily selected among those frequently deregulated in our transcriptomic analyses described above. In this setting, we could detect at least one positive TS/PS gene expression in 72% of the examined tumors ($n = 73$), whereas the corresponding control tissues were all negative ($n = 21$) (Fig. 1C; primers listed in table S6).

We hypothesized that the epigenetic alterations that characterize the cancer cells (3, 4) would likely drive the nonprogrammed activations of the TS/PS genes (12, 13).

This hypothesis was supported by monitoring the frequency of TS/PS gene activations in various cancers (our 1776 reference cancer samples) as a function of the status of their promoters, which demonstrated that the TS/PS genes associated with CpG-rich promoters, with high levels of methylation in somatic cells, were more frequently activated in cancers than genes associated with CpG-poor promoters (Fig. 2A).

This observation was further validated by analyzing the methylation-dependent transcriptional activity of TS/PS genes in human colon cancer cell line HCT116 bearing the inactivation of DNA methyltransferase 3b (DNMT3b) or the inactivation of both DNMT3b and DNMT1 [HCT116 DKO (double knockout)] (Fig. 2, B and C), which enabled us to identify genes activated by global DNA demethylation. Using our dedicated microarray, we showed that in HCT116 DKO, CpG-rich associated genes were also frequently activated, in contrast to genes with CpG-poor promoters, which did not respond to DNA demethylation (Fig. 2B). This was confirmed by qRT-PCR on 49 CpG-rich TS/PS genes, nearly half of which showed a clear activation in the DKO cells (Fig. 2C; primers listed in table S6).

This experiment also showed that, despite the genome-wide DNA demethylation in these cells, a significant number of CpG-rich TS/PS genes remain silent (Fig. 2C). This suggests that, although DNA demethylation may be necessary, in some cases, it is not sufficient to induce an ectopic gene expression.

The ectopic activation of 26 tissue-restricted genes in lung tumors is a strong and independent predictor of poor prognosis

Having found that the off-context expression of normally silent genes systematically occurs in cancer, we next investigated whether these genes could represent useful biomarkers by considering one cancer type, lung cancer. Lung cancer is one of the most common cancers in humans and is the most frequent cause of mortality by cancer in men and women (14). In the context of a clinical research program, we recruited a cohort of 293 lung cancer patients (recruited at Grenoble University Hospital, France), who received surgery. The group included 152 patients with early-stage cancer (T1N0) according to the TNM classification (tumor size, node positivity, and metastasis). For each of the study subjects, genome-wide transcriptomic analysis was performed on pretreatment diagnostic tumor samples, and pathological and clinical data were recorded, including overall and disease-free survival over a period of 5 to 10 years [described in table S7; full Affymetrix transcriptomic data available on the Gene Expression Omnibus (GEO) Web site under the reference GSE30219].

Applying the strategy described above, we could detect aberrant expressions of TS/PS genes in all lung tumor samples in our series

Fig. 1. Ectopic expressions of TS/PS genes detected in multiple cancers.

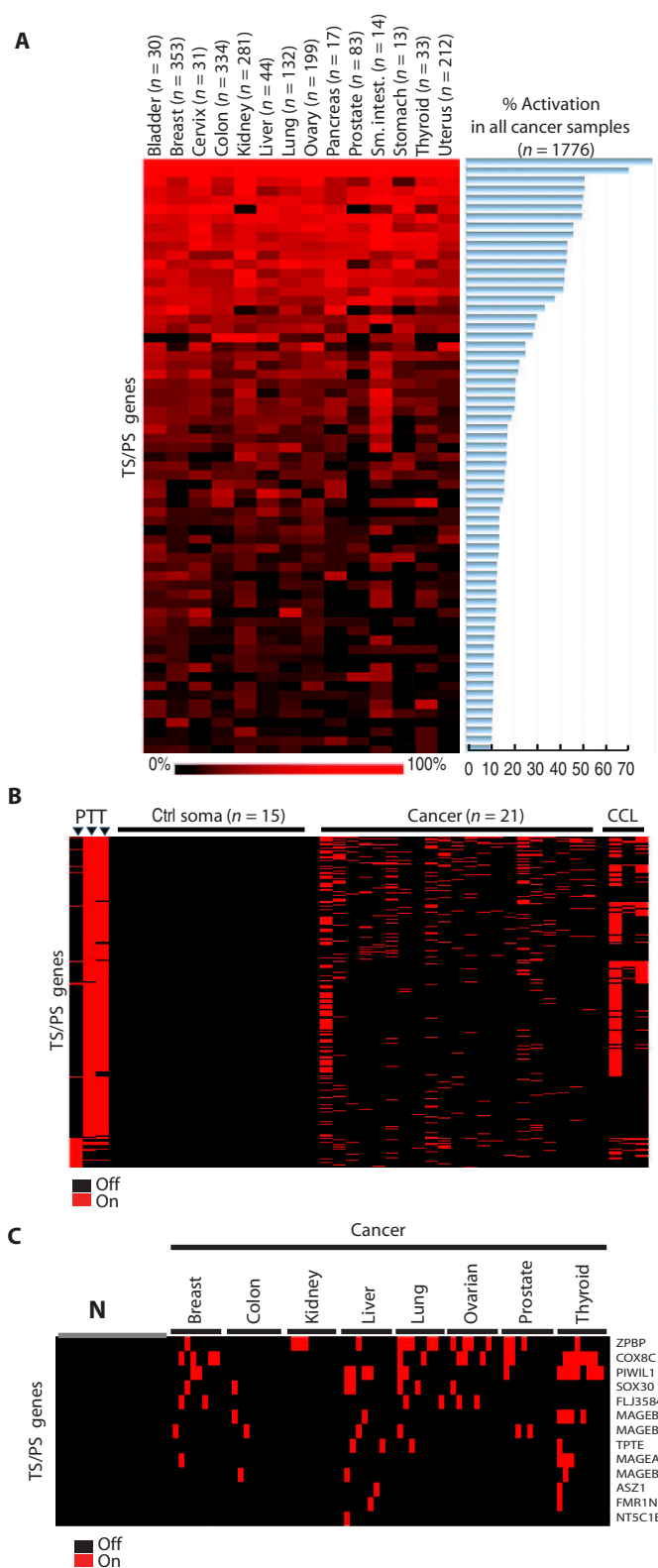
(A) Frequencies of activation of TS/PS genes in 14 different types of solid tumors from our analysis of transcriptomic data available online obtained from 1776 solid tumor samples (GSE2109, described in table S3) shown here for the 65 most frequently activated genes on a black (0%) to red (100%) scale (heat map). Overall frequencies of gene activation in all tumor samples are presented in the histogram on the right. **(B)** Transcriptomic profiles of TS/PS genes on a dedicated microarray in normal human tissues and in tumor samples and cell lines: "P," "T," "Ctrl soma," "Cancer," and "CCL" respectively indicate placenta, testis ($n = 2$), adult somatic tissues, cancer samples of various origins, and cancer cell lines ($n = 3$). All samples are listed in table S5. Color code: black, no expression ("Off"); red, gene activation ("On"). **(C)** Expression of a selection of 13 frequently expressed TS/PS genes detected by qRT-PCR in 73 tumor samples of eight different origins (including breast, colon, kidney, liver, lung, ovarian, prostate, and thyroid). The corresponding nontumor samples ("N") are shown on the left part of the heat map in the same order.

(Fig. 3A), including the 152 cases of early-stage T1N0 (Fig. 3B and table S8). Moreover, a series of 15 paired tumor and corresponding nontumor lung (NL) samples confirmed that these genes are mainly activated in the tumors and only rarely in the NLs of the same patients (these rare expression events in NLs could correspond to possible presence of residual cancer cells, or the effect of the tumor on the surrounding tissues) (Fig. 3C).

To look for epigenetic deregulations potentially leading to these ectopic expressions, we analyzed the methylome of a subset of 55 patients from the cohort. An initial analysis showed that the CpG-rich regions of the TS/PS genes, mostly hypermethylated in normal somatic tissues including lung, were globally demethylated in a large proportion of lung tumors (Fig. 3D and table S9). To explore a potential relationship between the demethylation of CpG-rich regions associated with TS/PS genes and their ectopic expression, we plotted the percentage of methylation of each CpG against the expression level of the corresponding gene (Fig. 3E and fig. S2). For most genes, ectopic activations were associated with the demethylation of at least one CpG, but, as suggested by our data from HCT116 DNMT DKO (Fig. 2C), the demethylation of any given CpG was not always associated with the activation of the corresponding gene. These observations, which agree with reported genome-wide cancer methylome studies (15, 16), suggest that, although DNA demethylation is associated with and involved in the ectopic gene activations in cancer, other factors/epigenetic aberrations also contribute.

We hypothesized that some of these aberrant expressions, by being the result of major epigenetic deregulations and/or leading to the expression of factors beneficial to cancer cells, could be associated with aggressive tumors and therefore have prognostic implications. Considering the TS/PS genes expressed in more than 1% of the tumors, we used univariate analysis to compare the global survival probabilities over a period of 5 years between the groups of patients whose cancers expressed each gene and those that did not.

This screen identified 26 TS/PS genes (fig. S1A, step 3), whose aberrant expression was individually associated with a lower survival probability in the lung cancer patients in our series [$P < 0.05$, log-rank test; hazard ratios (HRs) > 1.5 ; table S10]. To overcome the sporadic nature of these gene expression changes and optimize the information obtained on all samples, we decided to quantify the combined activations



of these 26 genes. We took a simple approach, which was as follows. The patients were first assigned into two groups: those with a tumor expressing none of the 26 genes, and those with a tumor expressing at least

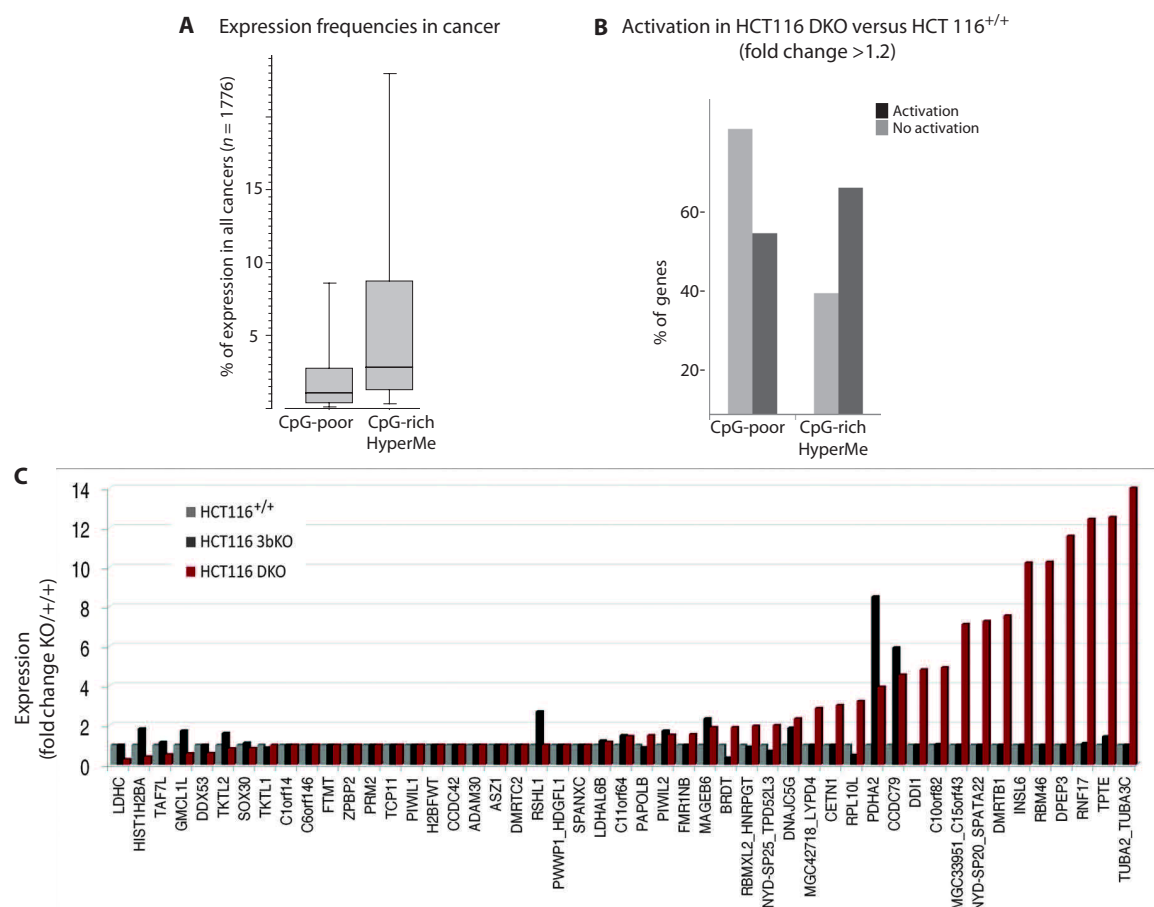


Fig. 2. TS/PS genes with hypermethylated promoters more susceptible to deregulation in cancer. (A) Box plots showing the respective distributions of activation frequencies in cancer (considering all 1776 cases of solid tumors analyzed in the data set GSE2109) of the two groups of genes, associated either with CpG-poor promoters (left) or with hypermethylated CpG-rich promoters (right). (B) Activation of TS/PS genes in response to DNA demethylation in HCT116 DNMT DKO (HCT116 cell line with double inactivation of DNMT1 and DNMT3b) compared to HCT116^{+/+} (wild type)

using our dedicated microarray; the histogram shows the numbers of genes activated (black) or not (gray) according to their promoter category, either CpG-poor or CpG-rich hypermethylated ("CpG-rich HyperMe"). (C) qRT-PCR detecting the expression of 49 TS/PS genes among those associated with a CpG-rich hypermethylated promoter (listed on the x axis of the histogram) in HCT116 cell line wild type (+/+), KO for DNMT3b (dark gray bars), and DKO for DNMT1 and DNMT3b (red bars); values are fold changes in reference to the normalized values obtained in HCT116^{+/+}.

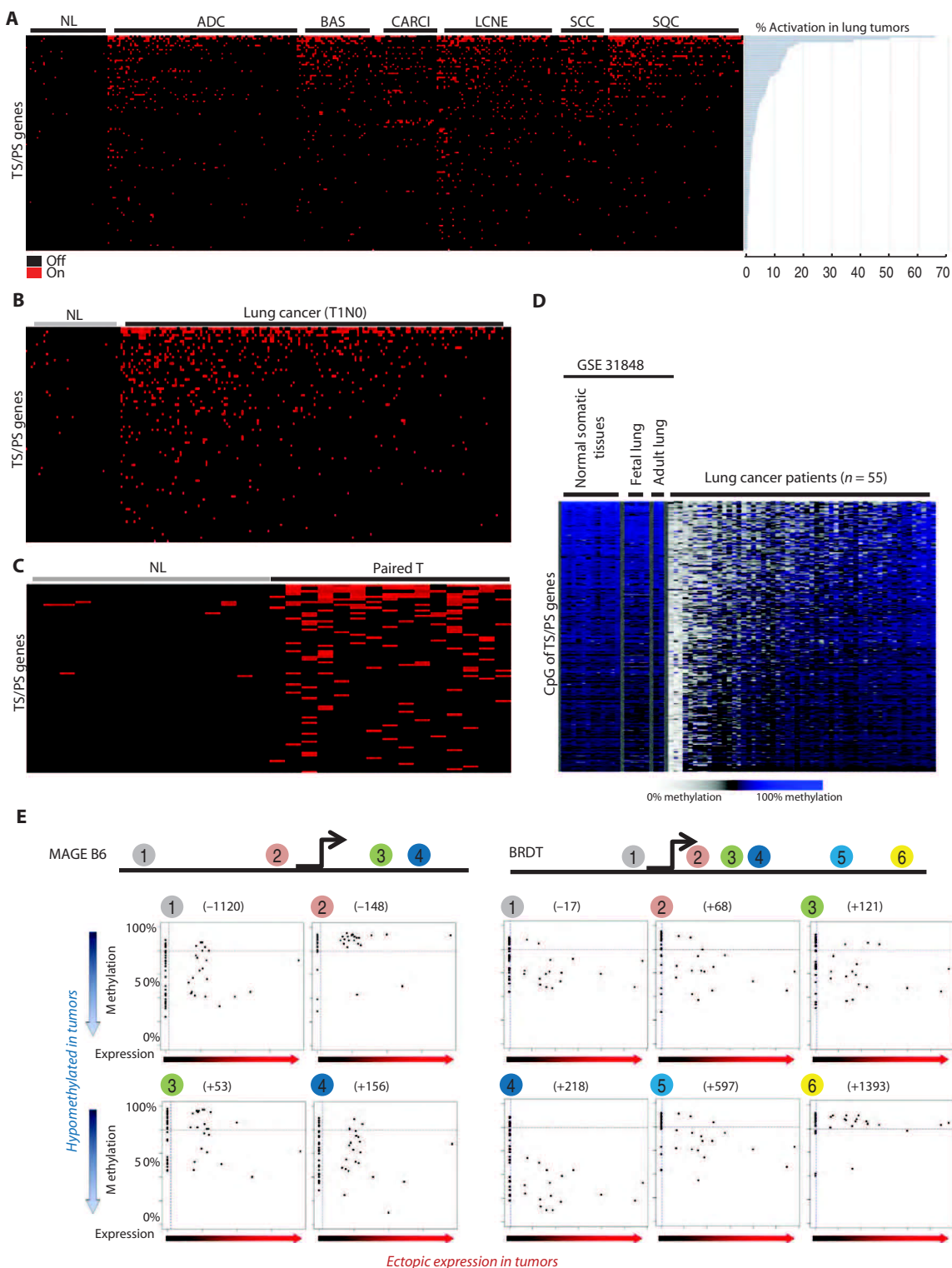
1 of the 26 genes. We then further refined this latter group by distinguishing tumors expressing one or two genes from tumors expressing three genes or more. As a result of this approach, the patients were stratified into three groups—P1, P2, and P3—according to the number of 26 genes expressed in their tumors: P1 tumors expressed none, P2 tumors expressed 1 or 2, and P3 tumors expressed 3 or more of the 26 genes.

We found highly significant differences in overall survival probabilities between these three groups ($P < 0.0001$; Fig. 4A). Additionally, the prognostic power of this 26-gene classifier was independent of other parameters, including clinical stage (TNM classification) (Fig. 4B) and histological subtype (fig. S3). In particular, this 26-gene group was a very efficient predictor for overall survival of early-stage patients. A multivariate analysis confirmed that our 26-gene combination was a stronger prognostic parameter associated with overall survival than histological subtypes or TNM stages ($P < 0.0001$; table S11). The survival predictive power of our 26-gene group was also strong when considering various subsets of patients grouped by sex, age, tobacco consumption, adjuvant chemotherapy or radiotherapy, as well as their

status for P53 mutation, as shown by the forest plot diagram of survival HRs between P1 and P3 patients (Fig. 4C). Finally, a comparison of the clinical outcomes between P1 and P3 patients allowed us to confirm that the tumors classified "P3" presented a particularly aggressive phenotype. Indeed, most patients with these tumors quickly relapsed and/or developed metastases, which was generally followed by short-term fatal outcome (Fig. 4, D and E).

To validate the potential of our genes to be used as a prognosis-classifying tool, we tested a subset of our patients using qRT-PCR detection of our genes. A clear-cut specific amplification was obtained for 25 of our classifying genes (Fig. 5A, left panel). A total of 61 patients in our series were tested using qRT-PCR and were assigned into P1, P2, and P3 groups, as defined previously. The comparison of survival probabilities between these three groups demonstrated our ability to use qRT-PCR to predict prognosis (Fig. 5A, right panel). A comparison between qRT-PCR and the transcriptomic approach for the detection of ectopic expressions of our prognostic genes showed that 44 of the 61 patients (72%) were assigned into similar groups by the

Fig. 3. Ectopic expression of TS/PS genes in the series of 293 lung cancer patients at all stages of the disease. (A) The heat map (left panel) shows the detection of TS/PS gene expression in all 293 patients (x axis) including adenocarcinoma (ADC), basaloid (BAS), carcinoid tumors (CARCI), large cell neuroendocrine tumors (LCNE), small cell carcinoma (SCC), and squamous cell carcinoma (SQC) histological subtypes, as well as in NL samples. Heat map color code: black, no activation; red, activation. The histogram (right panel) shows the frequency of lung cancer tumors (x axis, in %) aberrantly expressing each of the same TS/PS genes (y axis). (B) Heat map showing the expression of TS/PS genes focusing on the early lung cancer (T1N0) cases ($n = 152$) of the series, as described above; all patients are included in (A), but here they are sorted by decreasing number of ectopically expressed genes (not by histological subtypes). (C) Heat map focusing on the expression of TS/PS genes in a subset of 15 paired tumor samples ("Paired T") and their corresponding NL [these samples are also shown in the previous two heat maps in (A) and (B), but in different order because they are of different TNM stages and histology]; this figure shows only the genes activated in this subset of patients. (D) Heat map showing the methylation levels (β values from 0 to 100% on a light gray to blue color scale) of 347 CpGs associated with the transcription start site (TSS) and 5' untranslated regions of 88 TS/PS genes in normal somatic tissue samples (from left to right: mean methylation value in adipose tissue, adrenal gland, bladder, blood, brain, heart, lymph node, pancreas, skeletal muscle, spleen, stomach, ureter, and five fetal and two adult lung samples; data available on the GEO Web site under the



reference GSE31848), as well as in 55 lung tumor samples of our series. β Values are shown in table S9. (E) Scatter plots corresponding to the individual CpGs localized near the TSS (-1500 to $+1500$ base pairs) of two genes, BRDT and MAGEB6, showing that the methylation levels (β values on the y axis) correlated with their respective expression levels (\log_2 ratios) in the 55 lung cancer patients. The positions of the CpGs relative to the TSS are indicated between brackets.

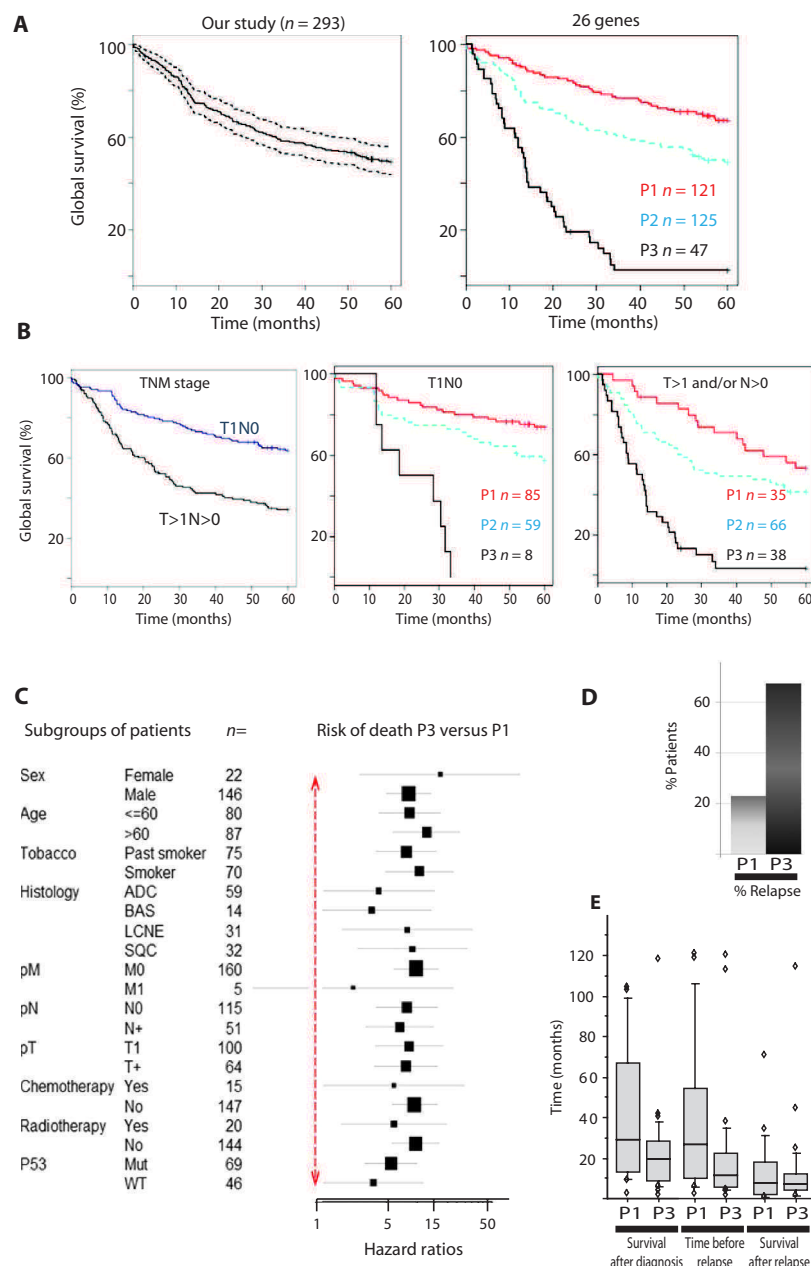


Fig. 4. Off-context activations of 26 TS/PS genes independently associated with poor prognosis in lung cancer. (A and B) Cumulative global Kaplan-Meier survival estimates of the 293 patients in our series either grouped together (A, left panel) or divided into three groups according to the number of ectopic expressions found within the subset of 26 genes. The groups were defined as follows: P1 (no expression, red curve, $n = 121$), P2 (one or two expressed genes, blue curve, $n = 125$), and P3 (three or more ectopically expressed genes, black curve, $n = 47$) (A, right panel). (B) The left panel shows the survival probabilities of patients according to the TNM stage (as indicated). The middle panel shows the survival probabilities of the three groups (P1, P2, and P3) defined by our classifying genes, considering only the T1N0 patients. The same approach was used to classify the T>1/N>0 patients (right panel). (C) Forest plot of HRs (P3 versus P1; on a log scale) for overall risk of death (more than 5 years). A univariate Cox proportional hazard model estimated HRs for the overall risk of death. The horizontal lines provide the 95% confidence interval for the ratios. The vertical red dotted double-arrow line corresponds to an HR of 1. (D) Histograms showing the frequencies of relapse (local recurrence and/or metastasis) observed in P1 and P3 patients. (E) Box plots showing the distribution of times (in months) for P1 and P3 patients, corresponding to overall survival (left), survival before relapse (middle), and survival after the diagnosis of relapse (right).

two approaches and that all of the patients in the poor-prognosis group P3 were correctly assigned using the qRT-PCR approach. A total of 17 patients were differently classified depending on the approach. Overall, these differences were due to the higher sensitivity of qPCR compared to the transcriptomic analysis. Indeed, five patients with short survival (<12 months) but assigned to the P1 group using the transcriptomic approach were better classified into the P2 to P3 groups by qRT-PCR, and three patients with intermediate survival (>12 and <36 months) were all assigned to the P1 group by the transcriptomic approach and to the P2 group by qRT-PCR detection. However, nine patients with survival of >36 months were more appropriately assigned to the P1 group using the transcriptomic approach than by qRT-PCR, which allocated them into the P2 group. An explanation for this observation is the “leaky” silence of our marker genes, which could be detected by qRT-PCR, because of the high sensitivity of this approach, leading to the misclassification of P1 tumors into the P2 group. Although these misclassifying expressions picked by qRT-PCR are not frequent, to remain “safe” in our prognosis classification, we decided to consider P2 tumors as an intermediate population where prognosis could not be accurately predicted.

This association between activations within our 26-gene group and a reduced global survival time was validated by applying the same 26-gene classification system to two external lung cancer populations, for whom survival follow-up times and transcriptomic data were obtained with the same Affymetrix technology as ours. These two studies included 82 patients (ADC, SQC, and large cell carcinomas) from (17) (GEO reference GSE19188) and 138 patients (ADC and SQC) from (18) (GEO reference GSE8894). According to the expression of our 26 classifying markers, these patients were assigned into P1, P2, and P3, and their survival times were compared. As mentioned above, because the P2 prognostic classification is not highly reliable for predicting prognosis, the corresponding survival curves are indicated as dotted lines. However, despite the very different origins of the two populations and the differences in the proportions of genders and histological subtypes, our 26-gene classifying system enabled us to reliably identify the very aggressive P3 type tumors in both studies (Fig. 5B), confirming the validity of our 26-gene classification in these independent external studies. In addition, the predictive power of a subset of these 26 genes was tested in a study of 443 patients with lung ADC (19), although, not using the full capacity of our classifier because only 11 of the 26 genes were represented on the Affymetrix platform used in this study, we could also confirm the discriminative power of our approach in this large and independent group (fig. S4).

Metastasis-prone aggressive lung tumors identified by the 26 genes have a characteristic molecular profile

The fact that our classifying genes could distinguish a group of very aggressive lung tumors (P3) independently of other

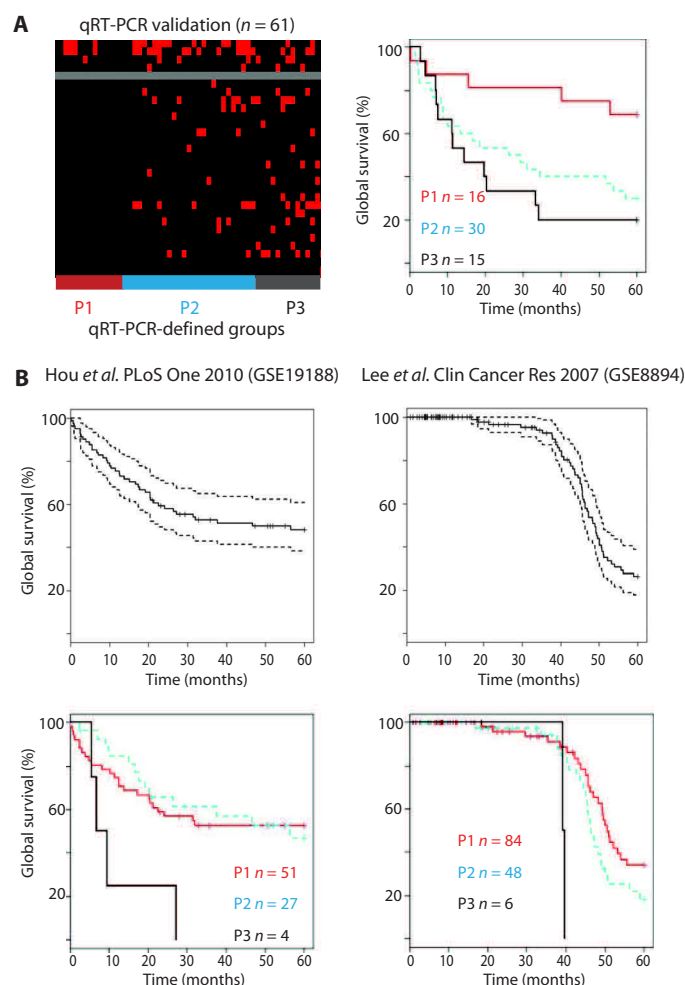


Fig. 5. Validation of the 26-gene prognosis-classifying strategy for lung cancer patients. (A) qRT-PCR detection of the expression of our classifying genes. The heat map (left) shows the detection of four frequently expressed TS/PS genes (upper panel) and of the prognosis-classifying genes (lower panel) in a subset of 61 patients from our lung cancer patient series. The survival curves (right) compare the survival probabilities between patients assigned to the P1, P2, and P3 groups by qRT-PCR. (B) Cumulative global Kaplan-Meier survival estimates of the patients from two external lung cancer populations either combined (top panels: the continuous lines show the mean survival probability, and the dotted lines correspond to the 95th percentile) or divided into the three groups P1, P2, and P3 defined as in Fig. 4A.

parameters prompted us to perform a supervised analysis of the transcriptome of these aggressive P3 tumors compared to the P1 group (fig. S1A, step 4) to identify, beyond our 26 classifying genes, gene expression patterns associated with this aggressive phenotype. This analysis revealed a gene expression profile composed of 1447 up-regulated and 1825 down-regulated genes (heat maps shown in Fig. 6A; genes listed in table S12), characterizing very aggressive tumors. A close inspection of the P3 versus P1 gene expression profile showed that a subset of patients that were ranked P1 (good prognostic) according to our 26 classifying genes presented a “P3-like” aggressive expression profile (Fig. 6A, blue rectangle). The survival of P1 patients with this P3-like profile was not significantly different from that of the other P1 patients, and differed from that of the

P3 patients (Fig. 6B), further demonstrating that the expression of the 26 classifying genes is a better prognostic predictor/censor of lung cancer survival than the P3 gene expression profile alone. This analysis also revealed another tumor group showing an intermediate gene expression profile: ones that did not overexpress P3 up-regulated genes but did present the P3 down-regulated gene profile (Fig. 6A, gray rectangle). All of these “intermediate” tumors corresponded to carcinoid tumors, which are associated with good prognosis.

On the basis of the expression profile of the highly aggressive P3 tumors, we attempted a biological characterization of these lung cancer “killer cells.” A GSEA enabled an in-depth investigation of the molecular profile of these aggressive tumors. The aggressive signature was highly enriched in genes normally predominantly expressed in embryonic stem or germline cells (Fig. 6C, upper panels). The analysis of the most enriched Gene Ontology (GO) terms showed that the up-regulated factors were mostly nuclear and related to the cell cycle and proliferation, whereas the depleted functions were related to the immune response and cell interactions and signaling (Fig. 6C, middle and lower panels). Accordingly, the analysis of the expression patterns of the down-regulated genes showed that 20% were genes that are normally predominantly expressed in immune organs (data obtained with normal samples of lymphocytes, spleen, or tonsil; listed in table S1), suggesting a reduced number of infiltrating immune cells within the tumors and/or in their environment. These data paint a molecular portrait of very aggressive and metastasis-prone cancer cells as those pushing their proliferative and self-renewal capacities while escaping immune surveillance systems and/or depleted in immune cells.

The aggressive lung tumor profile overlaps with signatures of other metastasis-prone tumors and could orient therapeutic strategies

The GSEA analysis also confirmed the shared characteristics between the aggressive signature of our series of lung cancer and the profile of the lung tumors associated with poor prognosis identified in the lung ADC study from (19) (Fig. 6D, left panels). The general transcriptomic profile corresponding to the aggressive tumors of this study largely overlapped that of the P3 tumor group we identified with the 26-gene signature, which further supported the relevance of our approach for the classification of the metastasis-prone lung tumors.

Other overlapping data sets confirmed our characterization of the P3 subset of lung tumors as an aggressive clinical phenotype, because we have also found very large overlaps with signatures from metastatic tumors of different origins (Fig. 6D, right panels, and table S13).

The expression profile of our aggressive lung tumors also suggested that they could be resistant to doxorubicin and gefitinib (Fig. 6E). Other significantly enriched gene sets suggested that the genes that are highly up-regulated in these aggressive tumors could be down-regulated by specific therapeutic approaches, including treatment with the Ras inhibitor salirasib or Aplidin, a marine-derived compound with potential anticancer properties (20). Indeed, transcriptomic data of cells treated with these compounds revealed that the P3 molecular profile could be shifted to a more P1-like situation after administering these drugs (Fig. 6F and table S13).

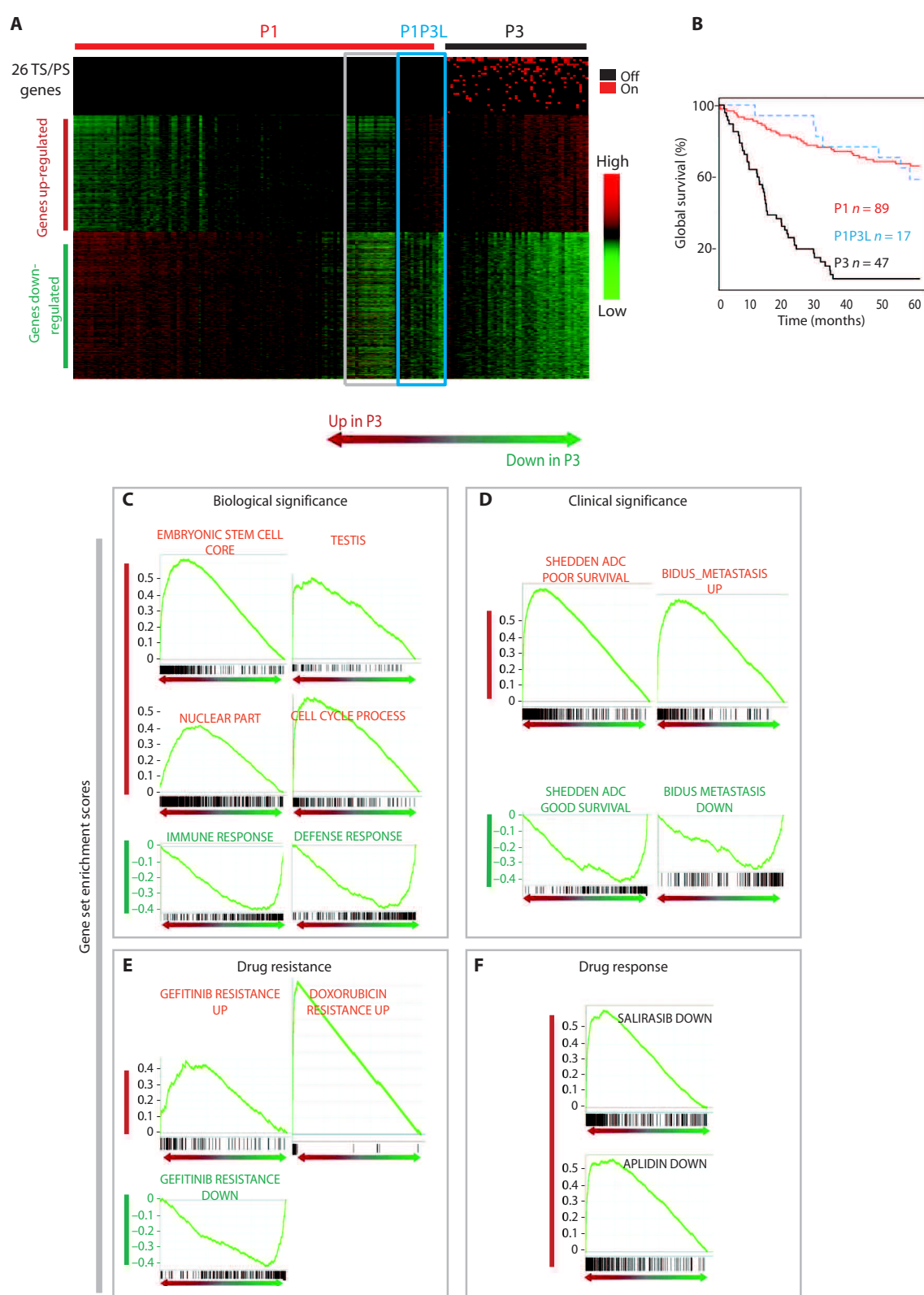
DISCUSSION

Here, we demonstrate systematic ectopic activation of hundreds of tissue-specific genes in many cancers, a phenomenon that had been

Fig. 6. Biological and clinical characteristics of aggressive lung tumors revealed by differential expression profiling. (A) Heat maps showing the expression of the 26 TS/PS genes (upper panel), as well as genes up-regulated (middle part) and down-regulated (lower part) in aggressive tumors (P3) compared with the “good prognostic” group of tumors (P1).

The robust multiarray average normalized values of expression of these genes are represented in the indicated color scales (green, low expression; red, high expression); the patients, represented on the x axis, were classified by prognostic groups and ranked by increasing value of the differential expression between up- and down-regulated genes; genes were ranked by decreasing difference of expression values between P1 and P3 samples. Blue frame: subset of the patients classified as group P1 presenting a P3-like molecular profile (P1P3L); gray frame: carcinoid tumors displaying an atypical molecular profile.

(B) The Kaplan-Meier curves represent the respective survival probabilities of the subgroups of P1 patients with the P3-like expression profile (P1P3L, blue curve), other P1 patients (red curve), and P3 patients (black curve). (C to F) Enrichment plots displaying the normalized enriched scores of some of the highly significant overlapping gene sets identified with gene set enrichment analysis (GSEA). The green curve shows the running enrichment score (y axis) for the gene set as the analysis walks down the ranked list of genes (x axis). The black bars along the x axis represent the genes of the gene set, ranked according to their fold change of expression in the “P3 versus P1” transcriptomic analysis (from left to right: up- to down-regulated genes).



previously reported to occur only sporadically. A key to this finding was the previous definition of tissue-restricted gene expression, which identified a specific category of genes that are normally expressed in the germline and placenta, and present a characteristic “locked” epigenetic configuration in all non-germline cell types, including embryonic stem cells.

The large-scale activation of this specific category of silenced genes in all cancers primarily reflects a general loss of cell identity, most likely due to a profound transformation of their overall epigenetic landscape. However, loss of methylation, although affecting most of the TS/PS genes promoters, cannot fully explain their ectopic activation. Changes in other epigenetic marks, such as histone modifications, should be considered to thoroughly evaluate the impact of the cancer epigenetic alterations in these aberrant gene activations.

Of particular interest is a group of 26 of these genes whose combined activation in lung tumors is associated with a particularly aggressive cancer phenotype.

The analysis of the literature on these 26 classifying genes showed that most of them had not previously been associated with cancers and that they have completely unknown functions.

This study therefore leaves us with important questions regarding the oncogenic potentials of these 26 genes. In particular, it is unclear how the ectopic activation of each of these genes or their combinations could account for the characteristic aggressive gene expression profile evidenced here in metastatic-prone aggressive tumors. It would also be important to understand the relationship between the expression of these genes and well-known oncogenic drivers.

An analysis of the literature shows that some of the 26 genes have already been identified as cancer-associated genes, mainly as so-called C/T antigens, and others are paralogs or pseudogenes corresponding to important cancer-related genes.

One of them, EB13 (Epstein-Barr virus-induced 3), has already been discovered not only as an independent predictor of poor prognosis in lung cancer but also as a driver (21). PIWIL1, also linked with poor prognosis in various cancer types including lung cancer, has been found associated with stem cell renewal (22–25), in agreement with the embryonic stem cell-like gene expression pattern observed in our aggressive P3 tumors. In addition, the identity of some of our genes suggests yet unknown oncogenic mechanisms. We also found two PTEN-related sequences among our 26 genes. TPTE encodes a testis-specific protein with high sequence similarity to the tumor suppressor PTEN (26), and TPTE2P2 encodes a testis-specific TPTE pseudogene. It is therefore possible that the ectopic activation of TPTE or a truncated protein produced by TPTE2P2 could interfere with the tumor suppressor activity of PTEN and act as a dominant negative factor with oncogenic activity. A similar mechanism could also explain the role of the NBPF4 neuroblastoma breakpoint family. Indeed, an amplified member of the family, NBPF23, has been recently shown to be specifically associated with pediatric neuroblastoma and could drive the malignant transformation through an unknown mechanism (27). The ectopic activation of NBPF4 could increase the NBPF gene dosage and elicit a similar oncogenic effect.

Our findings leave room for speculations and new investigations, but the very typical and homogeneous clinical, biological, and molecular portrait of aggressive tumors identified with this 26-gene signature implies that they might be associated with fundamental oncogenic mechanisms, which cooperate to generate such aggressive tumors.

Our approach also indicated several drugs that might efficiently target P3 lung tumors. Indeed, GSEA identified major overlaps be-

tween the sets of genes down-regulated by these drugs and the genes overexpressed in our P3 lung tumors. Some of these drugs, including salirasib (28) and Aplidin (29, 30), have been unsuccessfully tested against non-small cell lung cancer in phase 2 clinical trials. Our observation suggests that they might only be effective in a restricted number of tumors with a characteristic P3-like gene expression signature. The actual sensitivity of these tumor cells remains to be demonstrated in vitro as well as in vivo, including in patients, but the present data suggest potential therapeutic strategies to explore for these deadly tumors.

MATERIALS AND METHODS

Identification of genes with tissue-restricted expression using EST and transcriptomic data

Combining large-scale analysis of ESTs (31) ([http://www.ncbi.nlm.nih.gov/sites/entrez?db=unigene&cmd=search&term=9606\[taxid\]&AND+tissue\[restricted\]](http://www.ncbi.nlm.nih.gov/sites/entrez?db=unigene&cmd=search&term=9606[taxid]&AND+tissue[restricted])) and transcriptomic data on normal human tissues available online, we systematically looked for genes with a tissue-restricted pattern of expression. Briefly, only genes predominantly expressed in germline or placenta cells with no ESTs in other tissues were retained as restricted. The genes on this list underwent another consideration, which was the assessment of their promoter/epigenetic status. This led to the establishment of a list of 506 genes considered to have a restricted pattern of expression in germline and placenta.

Analysis of the epigenetic characteristics of the promoter regions of TS/PS genes in somatic cells

Publicly available pangenomic data were used to determine the epigenetic status of the promoter regions of the TS/PS genes. The list of selected studies and data processing is detailed in the Supplementary Materials and Methods.

Transcriptomic data analysis and statistics for normal tissues and cancer samples

Transcriptomic data from normal tissues and cancer samples were obtained with the Affymetrix technology “Affymetrix.GeneChip. HG-U133_Plus_2.” We used data available either on the GEO Web site (<http://www.ncbi.nlm.nih.gov/geo/>) or from our own lung tumor series (GSE30219). The data were analyzed as described in the Supplementary Materials and Methods.

Thresholds of expression for germline- and placenta-restricted genes were established as described in the Supplementary Materials and Methods.

Analysis of the expression of TS/PS genes on a dedicated microarray and by qRT-PCR

RNA of each of the various control, cancer, and cell line samples was extracted or purchased from Cytomix (see table S5 for information on the origin of the RNA samples) and then processed for hybridization or analyzed by qRT-PCR as described in the Supplementary Materials and Methods.

Analysis of the effect of DNA demethylation in inducing the expression of TS/PS genes in somatic cells using our dedicated microarray and qRT-PCR experiments

The objective was to compare the expression levels of TS/PS genes between HCT116 cell line with a DKO for DNMT3b and DNMT1 and its wild-type counterpart by analysis on our dedicated microarray

(described above; the values were expressed as fold changes over the threshold of signal obtained in normal somatic tissues; genes were considered activated when the normalized signal value was above the threshold value of 1.2) and qRT-PCR analysis (described above).

TS/PS gene DNA methylation analysis

Whole-genome DNA methylation was analyzed in 55 patients from our cohort with the Illumina Infinium HumanMethylation450 assay (32), as described in the Supplementary Materials and Methods.

Assessment of prognostic value of the off-context expression of TS/PS genes in lung cancer patients

The prognostic value of the expression of TS/PS genes in lung cancer was assessed in univariate and multivariate analyses as described in the Supplementary Materials and Methods.

Identification and characterization of the molecular profile of aggressive lung tumors

The supervised transcriptomic analysis carried out to identify the genes differentially expressed between patients with best prognosis (P1) and those with poorest prognosis (P3) and the GSEA of the gene expression profile characterizing aggressive tumors are described in the Supplementary Materials and Methods.

SUPPLEMENTARY MATERIALS

www.sciencetranslationalmedicine.org/cgi/content/full/5/186/186ra66/DC1

Materials and Methods

Fig. S1. Epigenetic “locking” of genes restricted to germline or placental expression.

Fig. S2. CpG methylation in the TSS regions of TS/PS genes.

Fig. S3. Prognostic value of ectopic expression of the 26 genes in lung cancer of specific histological subgroups.

Fig. S4. Prognostic value of ectopic expression of a subset of the 26 prognostic genes in an independent study of 443 lung ADCs.

Table S1. List of the normal tissue samples from GEO studies used for our analyses of gene expression profiles in normal tissues.

Table S2. List of genes with an expression restricted to male germ cells or placenta (TS/PS genes).

Table S3. Description and list of the 1776 cancer samples analyzed from the Exp0 cancer study (GSE2129).

Table S4. Percentage of positive expressions of TS/PS genes in the 14 different types of cancer analyzed from the Exp0 cancer study (GSE2129).

Table S5. List of tissue samples and cell lines analyzed on the dedicated microarray.

Table S6. List of primers used for qRT-PCR experiments.

Table S7. Key clinical and biological data for the 293 lung cancer patients in our study.

Table S8. Percentage of positive expressions of TS/PS genes in the different histological subtypes of our series of lung tumors.

Table S9. Methylation levels (β values) of the selected 347 CpG sites associated with 88 TS/PS genes in the 55 lung tumors in our series.

Table S10. List of 26 genes associated with poor survival in lung cancer and results of univariate analysis of association with prognosis in our lung cancer patients.

Table S11. HRs from the Cox regression model testing the association with prognosis in our set of 26 genes in a multivariate analysis, along with pathological data and TNM stages.

Table S12. List of the 3272 genes that are differentially expressed between patients classified by our 26 TS/PS genes in the good prognostic group (P1) and patients in the poor prognostic group (P3).

Table S13. List of relevant gene sets identified by GSEA from the Broad Institute Database corresponding to various GO terms (C5) and curated signatures (C2), including resistance to drugs or response to drugs, overlapping with our list of up- and down-regulated genes in P3 aggressive lung tumors.

REFERENCES AND NOTES

- J. Mellor, P. Dudek, D. Clynes, A glimpse into the epigenetic landscape of gene regulation. *Curr. Opin. Genet. Dev.* **18**, 116–122 (2008).
- C. T. Ong, V. G. Corces, Enhancer function: New insights into the regulation of tissue-specific gene expression. *Nat. Rev. Genet.* **12**, 283–293 (2011).
- M. Esteller, Cancer epigenomics: DNA methylomes and histone-modification maps. *Nat. Rev. Genet.* **8**, 286–298 (2007).
- P. A. Jones, S. B. Baylin, The epigenomics of cancer. *Cell* **128**, 683–692 (2007).
- M. Berdasco, M. Esteller, Aberrant epigenetic landscape in cancer: How cellular identity goes awry. *Dev. Cell* **19**, 698–711 (2010).
- S. A. Bert, M. D. Robinson, D. Strbenac, A. L. Statham, J. Z. Song, T. Hulf, R. L. Sutherland, M. W. Coolen, C. Stirzaker, S. J. Clark, Regional activation of the cancer genome by long-range epigenetic remodeling. *Cancer Cell* **23**, 9–22 (2013).
- A. Radu, C. Pichon, P. Camparo, M. Antoine, Y. Allory, A. Couvelard, G. Fromont, M. T. Hai, N. Ghinea, Expression of follicle-stimulating hormone receptor in tumor blood vessels. *N. Engl. J. Med.* **363**, 1621–1630 (2010).
- A. J. Simpson, O. L. Caballero, A. Jungbluth, Y. T. Chen, L. J. Old, Cancer/testis antigens, gametogenesis and cancer. *Nat. Rev. Cancer* **5**, 615–625 (2005).
- M. Weber, I. Hellmann, M. B. Stadler, L. Ramos, S. Pääbo, M. Rebhan, D. Schübeler, Distribution, silencing potential and evolutionary impact of promoter DNA methylation in the human genome. *Nat. Genet.* **39**, 457–466 (2007).
- M. G. Guenther, S. S. Levine, L. A. Boyer, R. Jaenisch, R. A. Young, A chromatin landmark and transcription initiation at most promoters in human cells. *Cell* **130**, 77–88 (2007).
- A. Barski, S. Cuddapah, K. Cui, T. Y. Roh, D. E. Schones, Z. Wang, G. Wei, I. Chepelev, K. Zhao, High-resolution profiling of histone methylations in the human genome. *Cell* **129**, 823–837 (2007).
- S. Rousseaux, S. Khochbin, New hypotheses for large-scale epigenome alterations in somatic cancer cells: A role for male germ-cell-specific regulators. *Epigenomics* **1**, 153–161 (2009).
- S. N. Akers, K. Odunsi, A. R. Karpf, Regulation of cancer germline antigen gene expression: Implications for cancer immunotherapy. *Future Oncol.* **6**, 717–732 (2010).
- R. Siegel, D. Naishadham, A. Jemal, Cancer statistics, 2013. *CA Cancer J. Clin.* **63**, 11–30 (2013).
- F. Fang, S. Turcan, A. Rimner, A. Kaufman, D. Giri, L. G. Morris, R. Shen, V. Seshan, Q. Mo, A. Heguy, S. B. Baylin, N. Ahuja, A. Viale, J. Massague, L. Norton, L. T. Vahdat, M. E. Moynahan, T. A. Chan, Breast cancer methylomes establish an epigenomic foundation for metastasis. *Sci. Transl. Med.* **3**, 75ra25 (2011).
- S. A. Selamat, B. S. Chung, L. Girard, W. Zhang, Y. Zhang, M. Campan, K. D. Siegmund, M. N. Koss, J. A. Hagen, W. L. Lam, S. Lam, A. F. Gazdar, I. A. Laird-Offringa, Genome-scale analysis of DNA methylation in lung adenocarcinoma and integration with mRNA expression. *Genome Res.* **22**, 1197–1211 (2012).
- J. Hou, J. Aerts, B. den Hamer, W. van Ijcken, M. den Bakker, P. Riegman, C. van der Leest, P. van der Spek, J. A. Foekens, H. C. Hoogsteden, F. Grosveld, S. Philipsen, Gene expression-based classification of non-small cell lung carcinomas and survival prediction. *PLoS One* **5**, e10312 (2010).
- E. S. Lee, D. S. Son, S. H. Kim, J. Lee, J. Jo, J. Han, H. Kim, H. J. Lee, H. Y. Choi, Y. Jung, M. Park, Y. S. Lim, K. Kim, Y. Shim, B. C. Kim, K. Lee, N. Huh, C. Ko, K. Park, J. W. Lee, Y. S. Choi, J. Kim, Prediction of recurrence-free survival in postoperative non-small cell lung cancer patients by using an integrated model of clinical information and gene expression. *Clin. Cancer Res.* **14**, 7397–7404 (2008).
- Director's Challenge Consortium for the Molecular Classification of Lung Adenocarcinoma, K. Shedden, J. M. Taylor, S. A. Enkemann, M. S. Tsao, T. J. Yeatman, W. L. Gerald, S. Eschrich, I. Jurisica, T. J. Giordano, D. E. Misek, A. C. Chang, C. Q. Zhu, D. Strumpf, S. Hanash, F. A. Shepherd, K. Ding, L. Seymour, K. Naoki, N. Pennell, B. Weir, R. Verhaak, C. Ladd-Acosta, T. Golub, M. Gruidl, A. Sharma, J. Szoke, M. Zakowski, V. Rusch, M. Kris, A. Viale, N. Motoi, W. Travis, B. Conley, V. E. Seshan, M. Meyerson, R. Kuick, K. K. Dobbin, T. Lively, J. W. Jacobson, D. G. Beer, Gene expression-based survival prediction in lung adenocarcinoma: A multi-site, blinded validation study. *Nat. Med.* **14**, 822–827 (2008).
- H. Depenbrock, R. Peter, G. T. Faircloth, I. Manzanares, J. Jimeno, A. R. Hanauske, In vitro activity of aplidine, a new marine-derived anti-cancer compound, on freshly explanted clonogenic human tumour cells and haematopoietic precursor cells. *Br. J. Cancer* **78**, 739–744 (1998).
- R. Nishino, A. Takano, H. Oshita, N. Ishikawa, H. Akiyama, H. Ito, H. Nakayama, Y. Miyagi, E. Tsuchiya, N. Kohno, Y. Nakamura, Y. Daigo, Identification of Epstein-Barr virus-induced gene 3 as a novel serum and tissue biomarker and a therapeutic target for lung cancer. *Clin. Cancer Res.* **17**, 6272–6286 (2011).
- L. F. Grochola, T. Greither, H. Taubert, P. Möller, U. Knippschild, A. Udelnow, D. Henne-Bruns, P. Würfl, The stem cell-associated Hiwi gene in human adenocarcinoma of the pancreas: Expression and risk of tumour-related death. *Br. J. Cancer* **99**, 1083–1088 (2008).
- X. Liu, Y. Sun, J. Guo, H. Ma, J. Li, B. Dong, G. Jin, J. Zhang, J. Wu, L. Meng, C. Shou, Expression of hiwi gene in human gastric cancer was associated with proliferation of cancer cells. *Int. J. Cancer* **118**, 1922–1929 (2006).
- H. Taubert, T. Greither, D. Kaushal, P. Würfl, M. Bache, F. Bartel, A. Kehlen, C. Lautenschläger, L. Harris, K. Kraemer, A. Meye, M. Kappler, H. Schmidt, H. J. Holzhausen, S. Hauptmann, Expression of the stem cell self-renewal gene Hiwi and risk of tumour-related death in patients with soft-tissue sarcoma. *Oncogene* **26**, 1098–1100 (2007).

25. H. Taubert, P. Würfl, T. Greither, M. Kappler, M. Bache, F. Bartel, A. Kehlen, C. Lautenschläger, L. C. Harris, D. Kaushal, S. Füssel, A. Mey, A. Böhnke, H. Schmidt, H. J. Holzhausen, S. Hauptmann, Stem cell-associated genes are extremely poor prognostic factors for soft-tissue sarcoma patients. *Oncogene* **26**, 7170–7174 (2007).
26. C. Tapparel, A. Reymond, C. Girardet, L. Guillou, R. Lyle, C. Lamon, P. Hutter, S. E. Antonarakis, The *TPTE* gene family: Cellular expression, subcellular localization and alternative splicing. *Gene* **323**, 189–199 (2003).
27. S. J. Diskin, C. Hou, J. T. Glessner, E. F. Attiyeh, M. Laudenslager, K. Bosse, K. Cole, Y. P. Mossé, A. Wood, J. E. Lynch, K. Pecor, M. Diamond, C. Winter, K. Wang, C. Kim, E. A. Geiger, P. W. McGrady, A. I. Blakemore, W. B. London, T. H. Shaikh, J. Bradfield, S. F. Grant, H. Li, M. Devoto, E. R. Rappaport, H. Hakonarson, J. M. Maris, Copy number variation at 1q21.1 associated with neuroblastoma. *Nature* **459**, 987–991 (2009).
28. G. J. Riely, M. L. Johnson, C. Medina, N. A. Rizvi, V. A. Miller, M. G. Kris, M. C. Pietanza, C. G. Azzoli, L. M. Krug, W. Pao, M. S. Ginsberg, A phase II trial of salirasib in patients with lung adenocarcinomas with KRAS mutations. *J. Thorac. Oncol.* **6**, 1435–1437 (2011).
29. T. Eisen, N. Thatcher, S. Leyvraz, W. H. Miller Jr., F. Couture, P. Lorigan, F. Lüthi, D. Small, A. Tanovic, M. O'Brien, Phase II study of weekly plitidepsin as second-line therapy for small cell lung cancer. *Lung Cancer* **64**, 60–65 (2009).
30. C. Peschel, J. T. Hartmann, A. Schmitt, C. Bokemeyer, F. Schneller, U. Keilholz, D. Buchheidt, S. Millan, M. A. Izquierdo, R. D. Hofheinz, Phase II study of plitidepsin in pretreated patients with locally advanced or metastatic non-small cell lung cancer. *Lung Cancer* **60**, 374–380 (2008).
31. M. S. Boguski, T. M. Lowe, C. M. Tolstoshev, dbEST—Database for “expressed sequence tags”. *Nat. Genet.* **4**, 332–333 (1993).
32. J. Sandoval, H. Heyn, S. Moran, J. Serra-Musach, M. A. Pujana, M. Bibikova, M. Esteller, Validation of a DNA methylation microarray for 450,000 CpG sites in the human genome. *Epigenetics* **6**, 692–702 (2011).

Acknowledgments: We thank J. Kim and his collaborators from Sungkyunkwan University (Seoul, South Korea) for sharing the raw data of their published work. We acknowledge the efficient contribution of the Centre de Ressources Biologiques of Grenoble Hospital in collecting and managing the cancer samples. **Funding:** The work in S.K./S.R. and E.B.'s laboratories is

supported by INCa-DHOS, ANR “EpiSperm,” and “ARC Subvention libre” programs. A.D. has been fully supported by INCa-DHOS, AGIRDOM, and ANR grants. D.G.B. was supported by grant R01 CA154365. The clinical research on lung cancer was funded by PNES POU MON INCA 2005 and BIOMARKSCAN PHRC 2003. The transcriptomic analyses were performed thanks to the program “Carte d'Identité des Tumeurs” supported by the Ligue Nationale Contre Le Cancer. S.R. and S.K. are recipients of a “contrat d'interface” from INSERM-Aviesan. **Author contributions:** S.R. and S.K. conceived the whole project, designed the experiments and data analysis approaches, and coordinated authors' contributions. S.R., B.J., A.V., J.-F.T., and A.D. performed the statistical and data analyses. P.-Y.B. ensured surgical removal of tumor samples. C.B., D.M.-S., and H.N.-M. were involved in the recruitment, clinical assessment, and follow-up of the patients. E.B. and S.L. ensured the histopathological assessment of the tumors. P.H. performed the screen for P53 mutations; A.d.R. and J.L. double-checked all the statistical analyses and performed the lung cancer methylome analyses. D.G.B. contributed to the transcriptomic and survival analysis in the lung ADC validation study. C.B., D.M.-S., H.N.-M., E.B., and S.L. were involved in tumor bank constitution and management. A.-L.V. performed the validation experiments. S.R. and S.K. wrote the manuscript, which was read and approved by all co-authors. **Competing interests:** The following patent applications include results presented in the paper: PCT/EP2009/053809, PCT/EP2011/068375, and PCT/EP2011/068377. **Data and materials availability:** The transcriptomic data from our series of lung tumors have been deposited in the GEO database (<http://www.ncbi.nlm.nih.gov/geo/>) under accession number GSE30219.

Submitted 16 January 2013

Accepted 12 April 2013

Published 22 May 2013

10.1126/scitranslmed.3005723

Citation: S. Rousseaux, A. Debernardi, B. Jacquiau, A.-L. Vitte, A. Vesin, H. Nagy-Mignotte, D. Moro-Sibilot, P.-Y. Brichon, S. Lantuejoul, P. Hainaut, J. Laffaire, A. de Reyniès, D. G. Beer, J.-F. Timsit, C. Brambilla, E. Brambilla, S. Khochbin, Ectopic activation of germline and placental genes identifies aggressive metastasis-prone lung cancers. *Sci. Transl. Med.* **5**, 186ra66 (2013).

FORUM ORIGINAL RESEARCH COMMUNICATION

Receptor-Independent Ectopic Activity of *Prolactin* Predicts Aggressive Lung Tumors and Indicates HDACi-Based Therapeutic Strategies

Aurore Le Bescont,¹ Anne-Laure Vitte,¹ Alexandra Debernardi,¹ Sandrine Curtet,¹ Thierry Buchou,¹ Jessica Vayr,¹ Aurélien de Reyniès,² Akihiro Ito,³ Philippe Guardiola,⁴ Christian Brambilla,¹ Minoru Yoshida,³ Elisabeth Brambilla,¹ Sophie Rousseaux,¹ and Saadi Khochbin¹

Abstract

Aims: Ectopic activation of tissue-specific genes accompanies malignant transformation in many cancers. *Prolactin* (*PRL*) aberrant activation in lung cancer was investigated here to highlight its value as a biomarker. **Results:** *PRL* is ectopically activated in a subset of very aggressive lung tumors, associated with a rapid fatal outcome, in our cohort of 293 lung tumor patients and in an external independent series of patients. Surprisingly *PRL* receptor expression was not detected in the vast majority of *PRL*-expressing lung tumors. Additionally, the analysis of the *PRL* transcripts in lung tumors and cell lines revealed systematic truncations of their 5' regions, including the signal peptide-encoding portions. *PRL* expression was found to sustain cancer-specific gene expression circuits encompassing genes that are normally responsive to hypoxia. Interestingly, this analysis also indicated that histone deacetylase (HDAC) inhibitors could counteract the *PRL*-associated transcriptional activity. **Innovation and Conclusion:** Altogether, this work not only unravels a yet unknown oncogenic mechanism but also indicates that the specific category of *PRL*-expressing aggressive lung cancers could be particularly responsive to an HDAC inhibitor-based treatment. *Antioxid. Redox Signal.* 00, 000–000.

Introduction

THE DEMONSTRATION OF large-scale ectopic activations of tissue-restricted genes in many unrelated cancers (1, 4, 9, 19, 23, 30) indicates that cancer cell hallmarks could be sustained by largely unknown mechanisms involving not only genetic alterations but also tissue-specific genes that are wrongly activated (9, 24, 25, 31, 37). The out-of-context activity of these tissue-specific genes, aberrantly expressed due to genetic or epigenetic upheavals, could be highly oncogenic. A clear example is that of NUT, a testis-specific factor expressed in a very aggressive midline carcinoma known as NUT midline carcinoma (NMC) in fusion with BRD4, a double bromodomain-containing factor (12). The dissection of the molecular actions of NUT in this context demonstrated the ability of NUT to sequester CBP/p300 in nuclear foci, subsequently leading to the inability of cancer

Innovation

Diverse and tumor-specific genetic and epigenetic upheavals underlie the occurrence of many yet unknown oncogenic mechanisms. Part of these oncogenic events comprises ectopic activation of normally tissue-specific genes (30). One of these genes is *prolactin* (*PRL*) reported here to be illegitimately activated in very aggressive lung tumor and associated with particularly poor prognosis. In breast and prostate cancer, oncogenic *PRL* acts through its receptor (*PRL-R*) following an autocrine or paracrine pathway (16). In this study, we show that a combination of genetic alterations and ectopic activation of *PRL* is at the heart of a completely unexpected receptor-independent oncogenic mechanism that can be counteracted by histone deacetylase (HDAC) inhibitors.

¹INSERM U823, Université Joseph Fourier–Grenoble 1, Institut Albert Bonniot, Grenoble, France.

²Cartes d'Identité des Tumeurs Program, Ligue Nationale Contre le Cancer, Paris, France.

³Chemical Genetics Laboratory, RIKEN, Wako, Japan.

⁴INSERM U892, Centre de Recherche sur le Cancer Nantes Angers and UMR_S 892, Université d'Angers, Plateforme SNP, Transcriptome & Epigénomique, Centre Hospitalier Universitaire d'Angers, Angers, France.

cells to respond to various stimuli (29, 32). These molecular studies also indicated therapeutic strategies based on the histone deacetylase inhibitor (HDACi) (29, 32) or bromodomain inhibitors (11). In the case of lung cancer, a systematic and detailed investigation of ectopic gene activations showed that illegitimate expressions of a subset of germline and placenta-specific genes is specifically associated with highly aggressive tumors. This observation strongly supports the hypothesis that, among the hundreds of ectopically activated tissue-specific genes, some decisively contribute to the very aggressive nature of the cells. This hypothesis receives strong support from data obtained following the use of a brain tumor model induced in a specific *Drosophila* line. In this setting, tumor induction resulted in the ectopic activation of a panoply of germline genes that proved to be indispensable for sustaining tumor cell proliferation (21). Other examples show that in human cancers, individual germline-specific genes strongly contribute to the establishment/maintenance of the malignant states. For instance, various members of the MAGE family of cancer-testis genes have been shown to regulate the different signaling systems involving p53 (6, 27, 36, 38). Additionally, ectopic activation of germline-specific genes in cancers could also confer resistance to a specific anticancer treatment. An example is the illegitimate activation of *CYCLON*, a gene normally expressed in testis, in diffuse large B-cell lymphoma (DLBCL). *CYCLON* ectopic activation specifically interferes with the cell response to rituximab and is associated with poor clinical outcome in DLBCL (8).

Prolactin (PRL) and its receptor have now been convincingly shown to be involved in breast and prostate tumorigenesis (10), and consequently targeting the corresponding signaling systems could be considered as a reasonable therapeutic strategy in different cancers, especially in prostate cancer (16). To date, the tumor-promoting action of PRL has been thought to be dependent on its receptor (PRL-R), which is also expressed by cancer cells and enables them to respond, through paracrine or autocrine loops, to circulating PRL from the pituitary gland or to local PRL produced by the cancer cells themselves (10).

In this study, we report the investigation of the out-of-context expression of *PRL* in lung cancer to explore its re-

levance as a potential biomarker for prognosis and as a guide in orienting a particular therapeutic approach.

Results

Ectopic PRL activation in a subset of very aggressive lung tumors

Recently, we reported transcriptomic data from a cohort of 293 lung tumors (30). These data were used to monitor the expression of *PRL*, which was found to be expressed in 27 lung tumors out of our series of 293 (9.2%) (Fig. 1A, B). A search of expressed sequence tag (EST) database (Supplementary Fig. S1A; Supplementary Data are available online at www.liebertpub.com/ars), experimentally confirmed by RT-PCR and RT-qPCR approaches on RNA from various human tissues (Supplementary Fig. S1B, C), showed that *PRL* is normally highly expressed in the pituitary gland and also, at reduced levels, in several other tissues, including the placenta. This comprehensive investigation of *PRL* expression showed that, although *PRL* expression is not strictly restrained to the pituitary gland and can be detected in some other extrapituitary sources, no ESTs or transcripts were detected in normal lung (Supplementary Fig. S1A–C). To confirm the absence of the *PRL* transcript in normal lung, an RT-PCR-based detection of *PRL* mRNA was setup using the placenta as a positive control and three *PRL*+ [#1, 3, and 6] and three *PRL*– tumors [#2, 4, and 5]. The result obtained (Fig. 1C) along with the *in silico* studies described above definitively confirmed the absence of *PRL* mRNA in normal lung.

We became particularly interested in the functional consequences of the ectopic activation of *PRL* in lung tumors following the observation that most of the patients with *PRL*+ tumors died rapidly (*p*-value < 0.0001; Hazard ratio = 3.23) (Fig. 1D). This association between the ectopic *PRL* expression and adverse clinical outcome was confirmed in a completely independent study (transcriptomic data publically available under the GEO reference GSE19188), where the patients with *PRL*-expressing tumors (10.9%) also exhibited a very poor prognosis (Fig. 1E). Since neuroendocrine lung tumors are more prone to express neuronal genes (13, 35), we wondered whether the ectopic *PRL* activation could only occur in tumors of the neuroendocrine

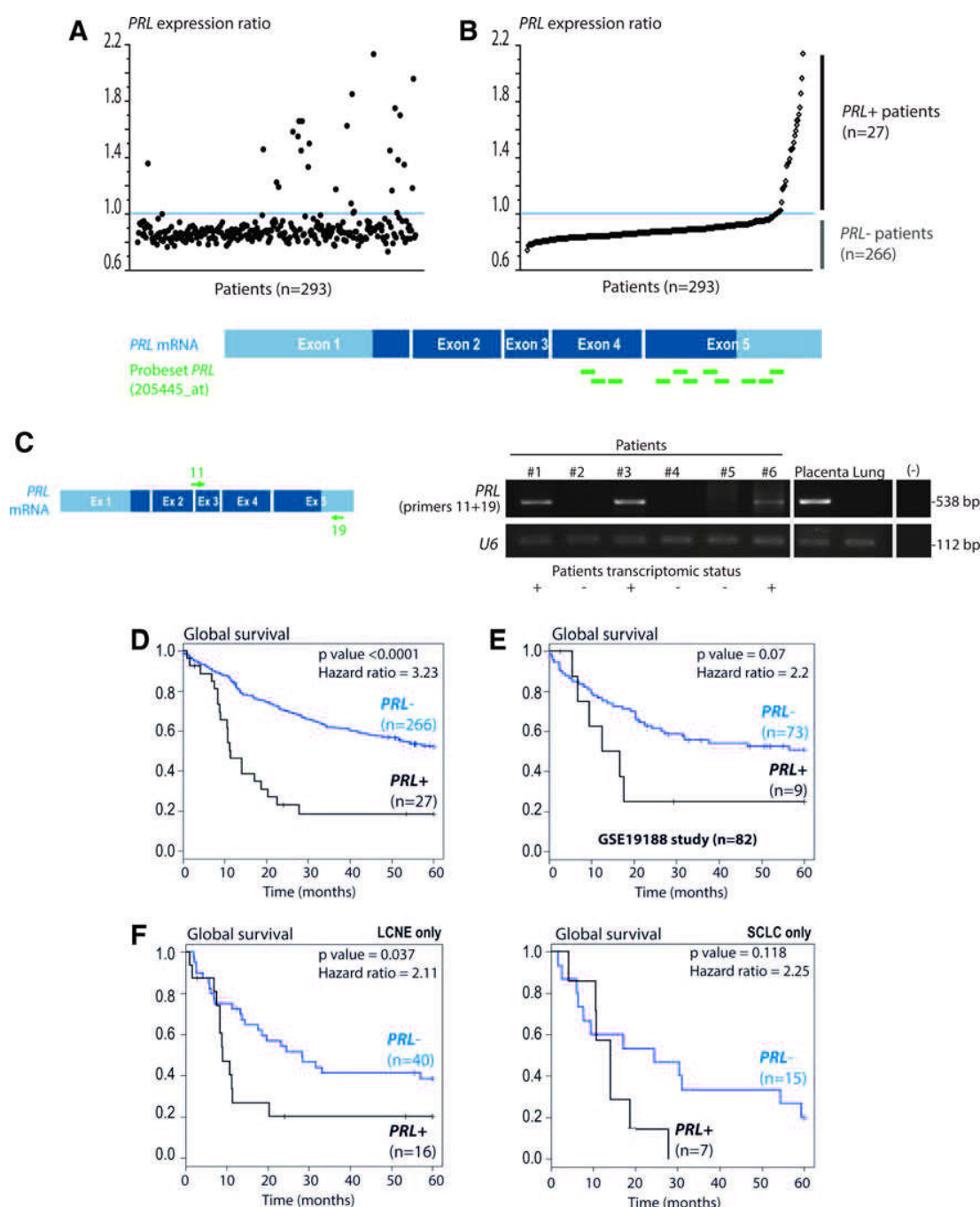
FIG. 1. *PRL* expression detected by a transcriptomic analysis in 9.2% of pulmonary tumors is associated with a patient's poor prognosis. (A, B) Expression ratios of *PRL* probe set (205455_at, 11 probes recognizing exons 4 and 5 of *PRL* gene sequence, whose positions are shown on the lower panel) obtained by transcriptomic analysis (Affymetrix Technology) of 293 lung tumors of the cohort. Expression ratios were obtained by dividing the RMA normalized signal values by the mean value of the *PRL* signal obtained from in 109 normal human somatic tissues (not including the pituitary gland). We considered that the *PRL* gene was expressed in a given tumor if this ratio was higher than the threshold (blue line: this threshold corresponds to the mean value of the *PRL* signal obtained from in 109 normal human somatic tissues as above). Twenty-seven tumors (9.2% of the cohort) exhibiting a *PRL* expression ratio higher than the threshold were considered *PRL*-expressing tumors; patients were labeled "*PRL*+ patients." (A) Patients were randomly assigned. (B) Patients were classified in an ascending order of *PRL* expression ratios. (C) RT-PCR analysis with *PRL*-specific primers of six patients' RNA samples. *U6* amplification was used as a loading control. Negative control (–) was performed with H₂O. The *PRL* status (+ : expressing or – : not expressing), as determined by the transcriptomic analysis, is indicated for each patient. The scheme on the left shows the position of the exons in *PRL* mRNA and those of the primers used (numbered). (D–F) Survival probability curves (established by the Kaplan–Meier method) for *PRL*– and *PRL*+ patients. The significance of the difference in prognosis was assessed by the univariate log-rank test. The hazard ratio estimates the difference in death rates between the two groups. (D) Survival probability curves in the lung cancer (CHU Grenoble) cohort (*n* = 293 patients with lung cancer) for *PRL*– (*n* = 266) and *PRL*+ patients (*n* = 27). (E) Validation performed on the GSE19188 external study (*n* = 82 patients with lung cancer). (F) Survival probability curves in the neuroendocrine group ectopically expressing (*PRL*+) or not (*PRL*–) *PRL*. LCNE (*n* = 56) (left) and SCLC (*n* = 22) (right) are shown. LCNE, large cell neuroendocrine; PRL, prolactin; SCLC, small cell lung cancer. To see this illustration in color, the reader is referred to the web version of this article at www.liebertpub.com/ars

subtype. The analysis of the lung tumors of our series demonstrated that the majority of the *PRL*-expressing tumors indeed belonged to the neuroendocrine types of LCNE or small cell lung cancers (SCLC). However, we also found that not all LCNE and SCLC tumors expressed *PRL* and that, within each histological subtype, the ectopic activation of *PRL* was associated with a particularly aggressive behavior of the tumors (Fig. 1F). These observations, together with a multivariate analysis quantifying the implication of various parameters (histological classification, tumor node-metastasis [TNM] parameters, *PRL* expression) in the survival (Supplementary Fig. S1D), suggested that *PRL* activation

was not simply a phenomenon associated with neuroendocrine lung tumors, but could itself be a strong predictor of poor prognosis.

Receptor-independent action of ectopically activated *PRL*

Since *PRL* is believed to act through a specific receptor, we wanted to know if the *PRL*-expressing lung tumor cells also expressed the corresponding receptor (*PRL-R*). The analysis of our series of 293 lung tumor transcriptomes showed that 59 tumors ectopically activated *PRL-R*. However, only 5 of



these tumors also expressed *PRL*, so that 22 out of the 27 *PRL*+ tumors did not show any expression of *PRL-R* (Fig. 2A). We were also able to demonstrate that *PRL-R* expression is not associated with tumor aggressiveness (Fig. 2B). According to the EST databanks (Supplementary Fig. S2A), and consistent with our RT-qPCR data (data not shown), the receptor is normally expressed in the breast, placenta, and uterus, but not in the lung.

These observations suggested a possible and previously undetected *PRL*-dependent, but receptor-independent oncogenic mechanism and prompted us to investigate the issue. Accordingly, to identify *PRL*+ cell lines, we screened the available transcriptomic data from lung cancer cell lines (data not shown) and from various cancer cell lines available to us (Fig. 3A). Following these approaches, we chose four cell lines, all from SCLC, three (H146, H524, and H69) expressing *PRL* at variable, but low amounts (not exceeding the placental expression level), and one (H526) showing no *PRL* expression (Fig. 3B, C). The analysis of the expression of *PRL-R* showed that none of our three *PRL*+ cell lines expressed the receptor (Fig. 3B). In this experiment, to make sure that we confidently defined the state of *PRL-R* expression, a PCR primer set was chosen in a region present in all *PRL-R* encoding mRNAs (2, 20) (Supplementary Fig. S2B) and confirmed the absence of *PRL-R* expression in the nonpathological lung and in our three *PRL*+ cell lines (Supplementary Fig. S2C). These cell lines appeared therefore as good model systems to investigate the molecular basis of *PRL* activity in lung cancer.

Ectopically expressed PRL mRNAs in lung cancers bear 5' truncations and systematically lack the signal peptide-encoding region

To better understand the possible receptor-independent cellular functions of the ectopically activated *PRL* in lung cancers, we first addressed the question whether an important functional determinant of the PRL protein, its signal peptide, was present in the ectopically expressed *PRL* mRNA. The absence or the presence of a nonfunctional signal peptide would lead to the production of a nonsecreted PRL and

the accumulation of the protein in the cells, which would have the potential to affect various cellular functions as opposed to a secreted protein that, in the absence of *PRL-R*, would not have any effects.

We designed a series of primers to better characterize *PRL* transcripts expressed in our *PRL*+ lung cell lines. We first focused on the two known transcript variants of *PRL*, differing only in the 5'UTR and encoding the same protein, which are expressed either from a proximal promoter in the pituitary gland (variant 1) or from a distal promoter used in the extrapituitary tissues (variant 2) (15). Although we could detect the expected transcripts in the pituitary gland and placenta, the transcripts detected in our lung cancer cell lines lacked the 5' regions, which are normally present in transcripts produced from either of the promoters (Fig. 4A).

A more detailed analysis of these transcripts with sets of primers spanning the *PRL* transcript from a fix 3' position toward the 5' end showed that in all the three *PRL*+ cell lines (H146, H524, and H69), *PRL* transcripts lacked its 5' exons (exon 1 and even a portion of exon 2) (Fig. 4B). Since the beginning of the signal peptide is encoded by exon 1, all the *PRL* transcripts produced in the three studied cell lines lack at least a portion of the signal peptide-encoding region.

Following this observation, we also tested 50 lung tumors, including 31 *PRL*+ and 19 *PRL*− samples. The 31 *PRL*+ group included 23 *PRL*+ tumors detected by the affymetrix transcriptomic analysis as well as 9 additional *PRL*+ tumors, which were identified by RT-PCR among the tumors originally classified as *PRL*− following affymetrix analysis, thanks to the sensitivity of PCR approaches.

Scanning *PRL* transcripts with some of the above-mentioned primers showed that 29 out of the 31 *PRL*+ tumors were missing regions encoding the signal peptide, with various extents of 5' deletions (Fig. 4C and Supplementary Fig. S3). These data therefore indicate that in lung tumors and cell lines, not only *PRL* is ectopically activated but also genetic alterations almost systematically affect the genomic region encoding the 5' part of the transcript that covers the signal peptide.

We also wondered whether the ectopic expression of *PRL* could be associated with amplification of the gene or its

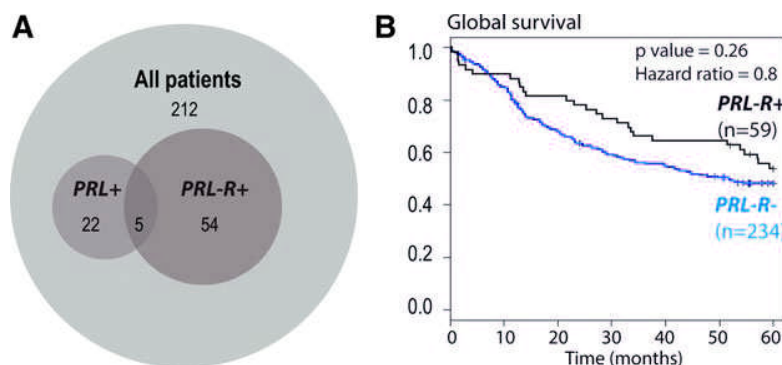


FIG. 2. Expression pattern of *PRL-R* gene in Grenoble lung tumors. (A) Expression of *PRL* and *PRL-R* genes in the 293 patients from the Grenoble lung tumors cohort. Expression values corresponding to the *PRL-R* probe set (227629_at) were extracted from the transcriptomic data (Affymetrix Technology) of 293 patients. The *PRL-R* gene was considered expressed if the *PRL-R* expression ratio was higher than a threshold corresponding to the mean signal for the *PRL-R* signal in 112 human somatic tissues + 2 standard deviations. *PRL-R* expressing tumors are labeled “*PRL-R*+.” (B) Survival probability curves (established by the Kaplan–Meier method) for *PRL-R*− ($n = 234$) and *PRL-R*+ patients ($n = 59$). The univariate log-rank test showed no significant differences between the two tumor groups. PRL-R, prolactin receptor. To see this illustration in color, the reader is referred to the web version of this article at www.liebertpub.com/ars

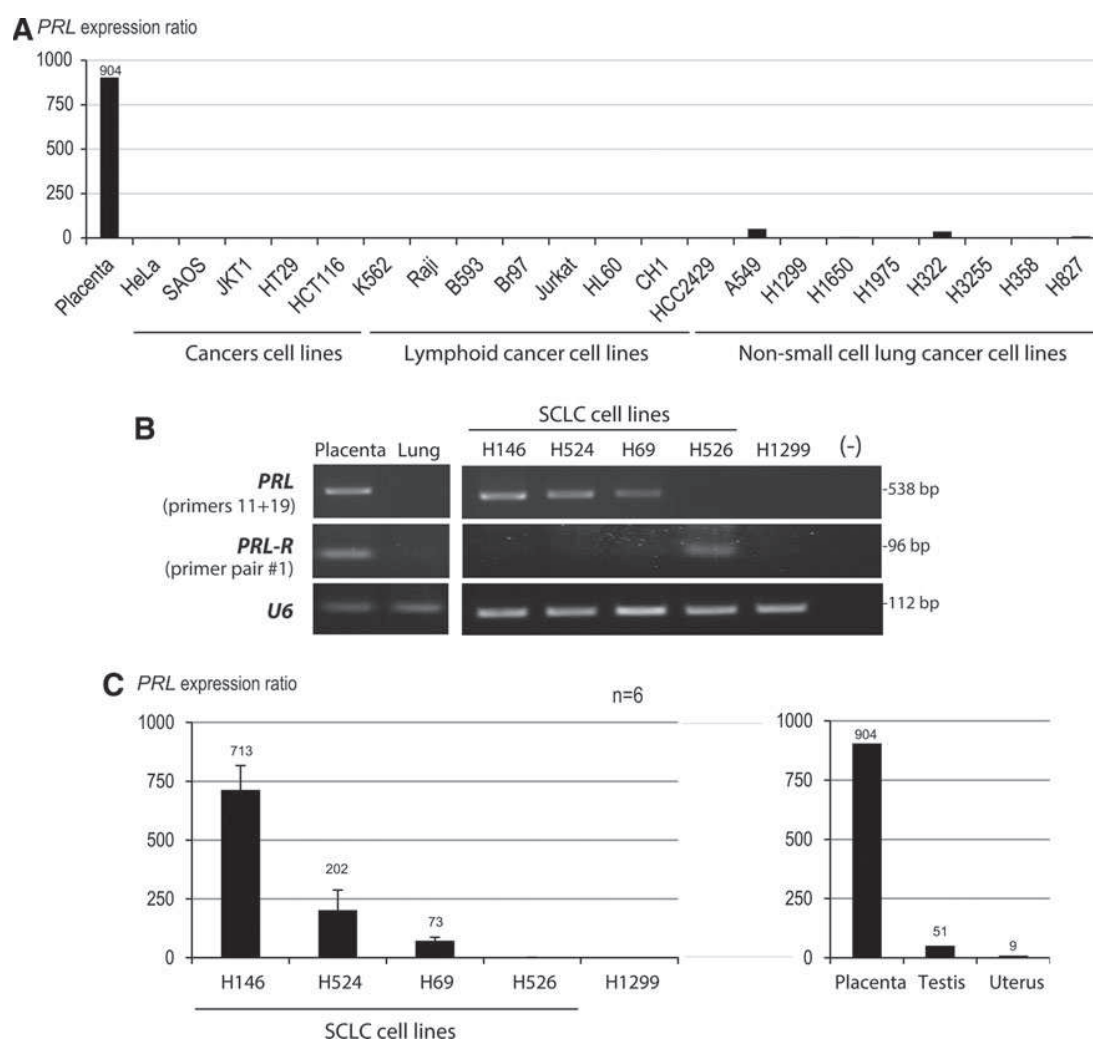


FIG. 3. Expression patterns of the *PRL* and *PRL-R* genes in human cancer cell lines: identification of SCLC cell lines expressing *PRL* and not expressing *PRL-R*. (A–C) RNA extracted from various human cancer cell lines were submitted to RT-PCR and RT-qPCR analyses with specific *PRL* primers. (A) RT-qPCR analysis on the indicated human cancer cell cDNAs. (B) RT-PCR analyses with primers specific for *PRL* and *PRL-R* on five human lung cancer cell cDNAs as indicated. (C) Expression ratios of the *PRL* gene in five human lung cancer cell lines (and comparison with tissues *PRL* expression ratios) obtained from six independent RT-qPCR experiments, with normalization using two different control genes (*Actin* and *U6*). One hundred percent indicates the mean value of *PRL* expression ratios from the six *PRL*-expressing tissues. Error bars represent the standard deviation.

genomic region. Data from comparative genomic hybridization (CGH) arrays, obtained for 247 of our 293 lung cancer patients, demonstrated that, although the genomic region corresponding to the *PRL* gene was amplified in a proportion of patients, there was no correlation between amplification and expression (Fig. 4D). Moreover, survival comparisons between groups of patients according to the presence or absence of amplification within this region demonstrated that there was no association between the amplification of this region and prognostic (data not shown).

*The ectopically expressed truncated *PRL* is a potent regulator of gene expression*

Our data show that all the *PRL*+ lung tumor samples and cell lines express a truncated form of a *PRL* transcript lacking

the 5' part of the mRNA to various extents. Since this almost systematic 5' deletion includes the translation initiation ATG codon, it is unlikely that *PRL* is produced. Nevertheless, we decided to investigate the consequences of the ectopic expression of these transcripts. This was achieved through a siRNA knockdown in two *PRL*+ cell lines (H146 and H524) followed by a whole genome transcriptomic analysis. With a threshold of 1.2 for mRNA variations between control siRNA and the anti-*PRL* siRNA, the number of genes affected was found to be proportional to the level of *PRL* expression. In the H146 cell line, expressing the highest level of *PRL* mRNA (Figs. 3C and 6A), *PRL* silencing resulted in the upregulation and downregulation of 2559 and 2402 genes, respectively, whereas in the H524 cells, which express about half the level of *PRL* mRNA (Figs. 3C and 6A), 800 and 763 genes were, respectively, upregulated and downregulated (Fig. 5A).

Interestingly, the majority of the genes affected by *PRL* silencing in the H524 cell line were also *PRL*-dependent genes in the H146 cells. Indeed, 74% and 68% of genes, respectively, upregulated and downregulated in H524 cells were included in the corresponding gene lists from the H146 cell line after *PRL* knockdown (Fig. 5A). These data strongly suggested that *PRL* affects the same regulatory pathways in these two different cell lines and that the extent of these effects depends on the amount of *PRL* expressed. Within the genes affected in both cell lines, we chose 12 for validation by RT-qPCR using H146 RNA from cells treated with control or anti-*PRL* siRNAs, and the data confirmed the robustness of the analysis (Fig. 5B). Hence, these results reveal an important gene regulatory function for these truncated *PRL* RNAs in both cell lines.

In an attempt to identify the cellular regulatory circuits controlled by *PRL*, a Gene Set Enrichment Analysis (GSEA) was performed on the list of genes differentially regulated after *PRL* knockdown in the H146 cell line. This analysis showed that the group of genes requiring *PRL* to be expressed (downregulated after *PRL* knockdown) significantly overlaps with previously established lists of genes upregulated in non-small cell lung tumors as well as in a variety of unrelated cancers (Fig. 5C).

Interestingly, the GSEA also revealed that in these cells, which are here maintained under normoxic culture conditions, the ectopic *PRL* expression activates a group of genes that were found to be also expressed in response to both hypoxia and HIF1 α (7) (Fig. 5C).

Remarkably, this analysis also revealed a highly significant overlap with genes that are upregulated either after treatment with the HDACi trichostatin A (TSA) or after the knockdown of HDAC1, HDAC2, or HDAC3 (18, 33) in independent studies and cell systems (Fig. 5C). These observations suggested that in *PRL*+ lung cancers, *PRL* somehow controls cellular regulatory circuits involving histone acetylation-dependent transcriptional activities, since both *PRL* knockdown and HDAC1–3 inhibition or knockdown affect a significant number of common genes. To better understand the relationship between *PRL* transcripts/peptides and HDACs, we treated our three *PRL*+ and one *PRL*– cell lines with TSA, as well as with the FDA approved HDAC

inhibitor, SAHA, and measured the level of *PRL* mRNA by RT-qPCR. Interestingly, and consistent with the GSEA results, we observed that treatment with both TSA and SAHA leads to a remarkable decrease in the steady-state amounts of *PRL* mRNA (Fig. 5D), indicating that by downregulating *PRL*, a treatment by HDAC inhibitors could also affect a significant number of common genes.

PRL downregulation sensitizes lung cancer cells to genotoxic treatments

Since in many cases, cancer cell treatment with HDAC inhibitors restores cell cycle check points and apoptotic cell responses (22, 26), we wondered if *PRL* downregulation could also render cells more sensitive to a cytotoxic treatment. Our three *PRL*+ and one *PRL*– lung cancer cell lines were treated with either a control or the two anti-*PRL* siRNAs (Supplementary Table S1A), leading to a significant knockdown of *PRL* expression in all *PRL*+ cell lines (Fig. 6A).

We then assessed the effect of *PRL* silencing on cell apoptosis. We first observed that the loss of *PRL* expression does not affect the basal mortality. We then subjected SCLC cells to genotoxic stress, through a treatment with actinomycin D (Acti D). In the three *PRL*+ cell lines (H146, H524, and H69), *PRL* knockdown increased the proportion of apoptotic cells following the Acti D treatment, while the *PRL* siRNA transfection in the *PRL*– cell line did not significantly modify the proportion of apoptotic cells. The latter cells appeared to be more responsive to a genotoxic treatment, which is consistent with the absence of *PRL* expression (Fig. 6B). This figure also shows the original FACS data for one experiment with the H524 cell line treated with control or anti-*PRL* siRNA (siRNA *PRL* #1) (Fig. 6C). These data show that *PRL* downregulation in this cell line under genotoxic treatment almost doubled the number of dying cells, which approached the totality of the population. It is of note that the dosage of Acti D used to induce genotoxic effects is below the amount required to block the *PRL* gene activity (Supplementary Fig. S4).

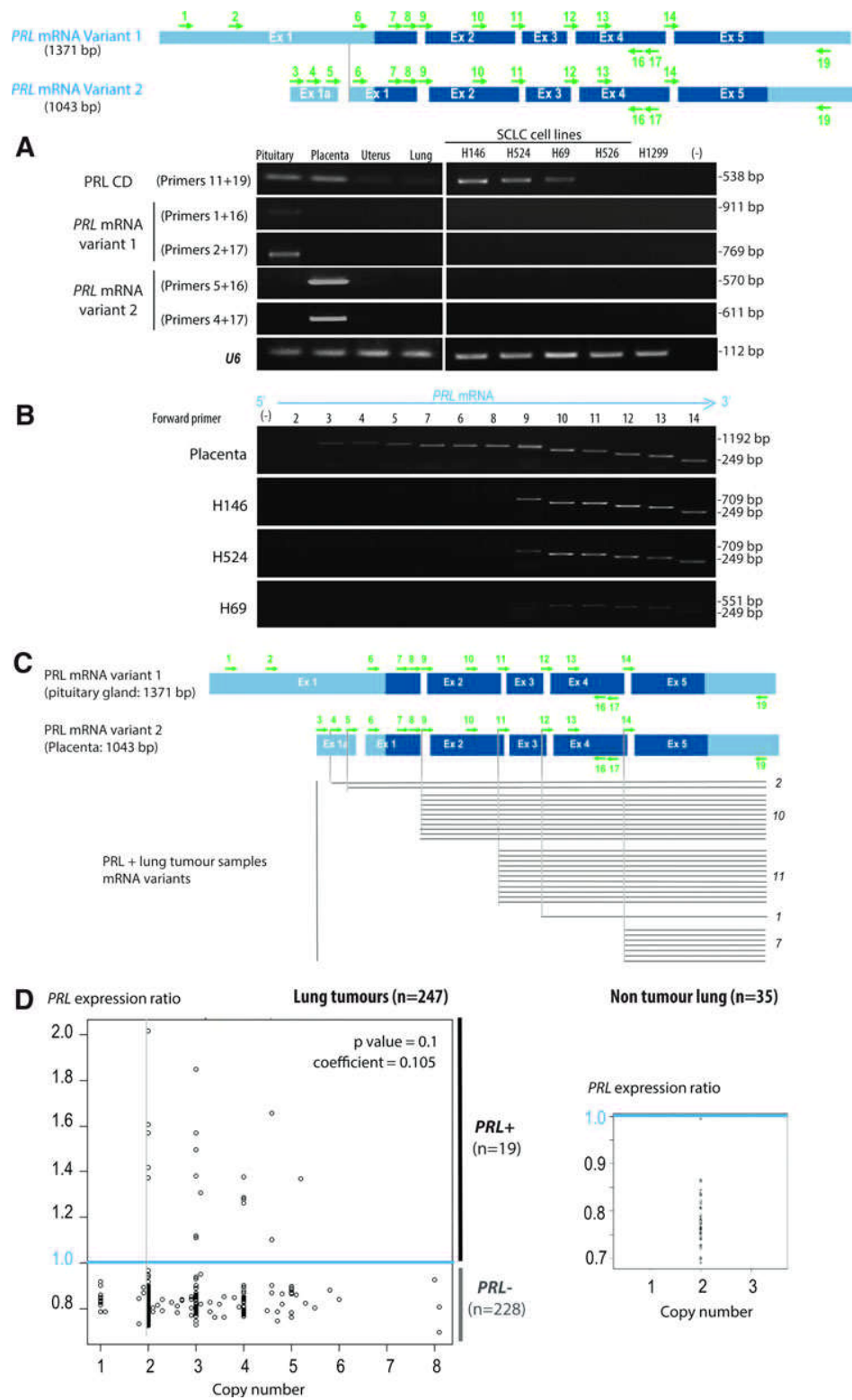
We concluded that the *PRL* ectopic expression in SCLC cells is involved in their resistance to cell death following a genotoxic treatment.

FIG. 4. The *PRL* mRNAs ectopically expressed in lung cancer cell lines or in tumor samples are lacking the 5' region corresponding to the signal peptide, but are not correlated with genomic amplification of the *PRL* gene region. (A) RT-qPCR analysis of *PRL* mRNAs from lung tumor cell lines. The *top panel* shows the position of the primers (numbered) used in (A–C). Primers 11, 16, and 17 are located within the *PRL* coding sequence (CD). Primers 1 and 2 are located within the 5'UTR sequence specific of the pituitary *PRL* mRNA (variant 1). Primers 4 and 5 are located in the exon 1a sequence, which is the 5'UTR sequence specific of the extrapituitary *PRL* mRNA (variant 2). RNAs extracted from human tissues and cancer lung cell lines were submitted to RT-PCR, with the indicated primer pairs. *U6* amplification was used as a control. Negative control (–) was performed with H₂O. (B) Cartography of *PRL* mRNAs shows that exon 1 and exon 1a sequences are absent from the transcripts from the SCLC cell lines. Cartography of the *PRL* mRNA expressed in the SCLC *PRL*+ cell lines was performed by using a reverse primer in 3'UTR (#19) and the indicated forward primers along the *PRL* transcript sequences (numbered). Negative control (–) was performed with H₂O. (C) Cartography of *PRL* mRNAs in tumor samples shows the extended absence of 5' sequences. Cartography of the *PRL* mRNAs generated in lung tumors was performed by using a reverse primer in 3'UTR (#19) and five different forward primers corresponding to every exon junction sequence (numbered). *U6* amplification was used as a control. Negative control (–) was performed with H₂O. (D) Analysis of variations in copy number (deletion and amplification) at the *PRL* gene locus by CGH array in 247 lung tumors of the cohort. For each tumor, the copy number at the *PRL* gene locus is indicated on the x-axis and the corresponding *PRL* expression value is indicated on the y-axis. The correlation between *PRL* expression and *PRL* copy number is evaluated by the nonparametric Spearman's coefficient. A control analysis on nontumor lung samples from 35 patients (*right panel*) is also shown. CGH, comparative genomic hybridization; UTR, untranslated region. To see this illustration in color, the reader is referred to the web version of this article at www.liebertpub.com/ars

Discussion

Genome instability and subsequent gross and subtle genome alterations associated with a profoundly modified epigenetic signaling, systematically accompany malignant

cell transformation (34). Growing and invading cancer cells harbor new states of gene expression involving not only activation and repression of normal genes but also *de novo* creation of genes that never existed in normal cells. Such genes result from chromosomal translocations, mutations,



loss and gain of specific genomic regions. The combined action of these created new functional factors, with gene silencing and ectopic activations of hundreds of tissue-restricted genes, gives birth to cells that acquire new properties (9, 24, 31) allowing them to sustain growth and dissemination, while avoiding destruction (17). In this study, we show that *PRL*, a gene never expressed in normal lung cells, becomes ectopically activated in lung cancer tumors, essentially in a fraction of the neuroendocrine subtype. This activation of *PRL* is strongly associated with adverse clinical outcomes, not only in our population of lung cancer patients from the Grenoble university hospital but also in a completely unrelated group of patients. Unexpectedly, we also found that, in both groups of patients, the majority of *PRL*+ lung tumors did not express detectable amounts of *PRL-R* transcripts. However, our present knowledge of *PRL* biology suggested that, in the absence of a receptor, the produced *PRL* should have no effects on the *PRL*-expressing tumor cells, since most of the synthesized proteins would be secreted out of the cells and no signaling could be initiated in the absence of the receptor. This was inconsistent with our clinical data, showing a systematic association between *PRL* expression and rapid fatal outcome, which suggested that the produced *PRL* should have an effect on the cells and somehow promote their aggressive behavior. Therefore, we made the hypothesis that, in the absence of its receptor, the only way *PRL* expression could have an impact on cells would be that the protein somehow would lose its signal peptide, and hence becomes trapped inside the cells. Our investigations in the selected *PRL*+ cell line models and tumors revealed genome-altering events, which almost systematically affected the first exon of *PRL*. In fact, in most of the tested cases, the transcript did not contain the translation initiation codon ATG. Even in the five cases of tumors expressing *PRL-R* and *PRL* mRNA, the *PRL* transcripts were truncated.

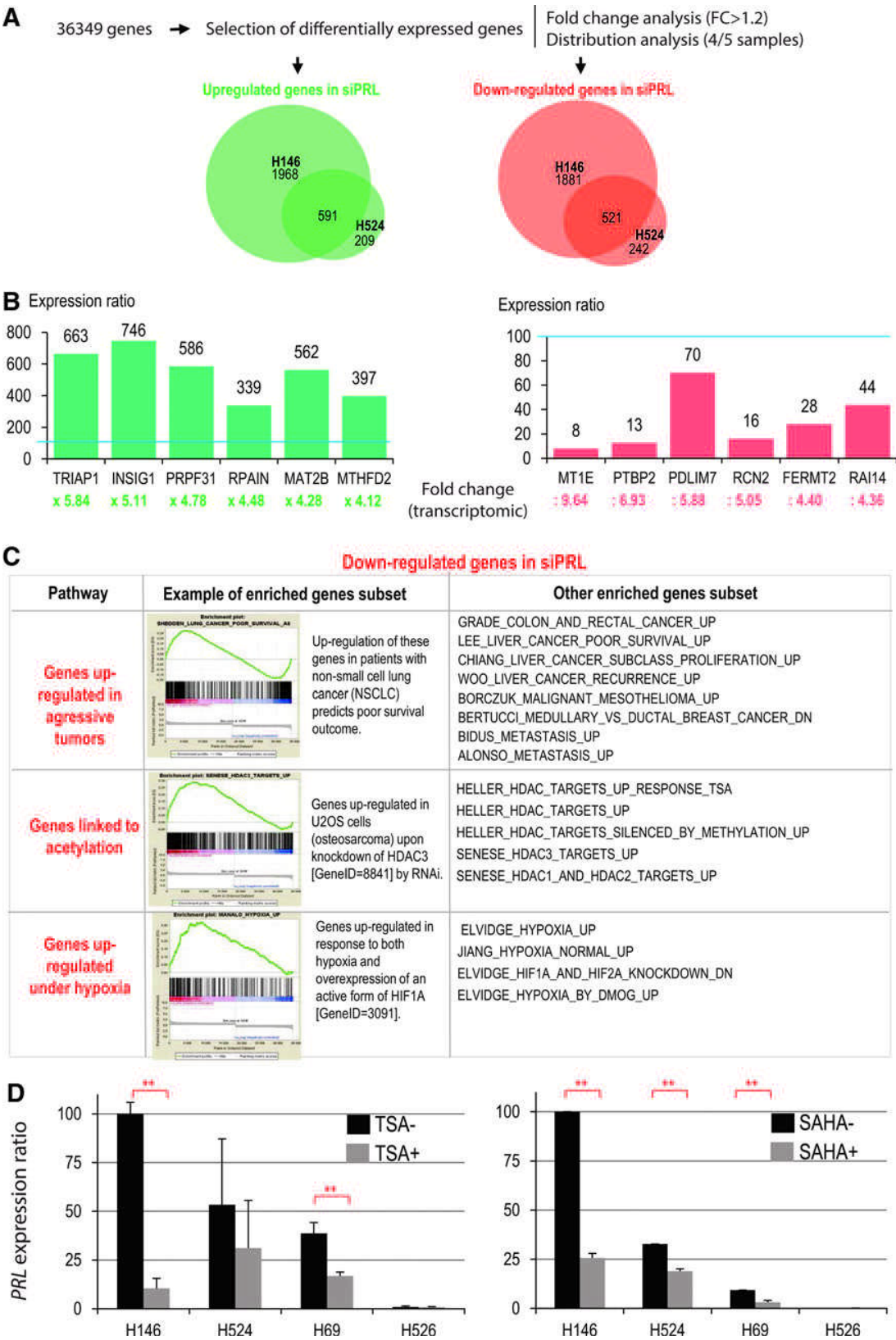
The important question following this observation is whether these mRNAs could give rise to a polypeptide. The use of an antibody able to detect large amounts of

PRL in the pituitary gland, did not allow us to visualize the protein neither in lung tumors nor in our lung cancer cell lines (data not shown). Therefore, at this point, the absence of the translation initiation codon ATG in the transcripts, suggests that actually no easily detectable peptides are expressed. Nevertheless, the detailed investigation of our *PRL*+ cells revealed clear and characteristic molecular responses to the knockdown of *PRL* mRNA. Most remarkably, *PRL* mRNA silencing was associated with very consistent alterations in gene expression in two different cell lines. This gene expression response was proportional to the level of *PRL* expression: a higher number of genes were affected in the cell line expressing the highest amount of *PRL* mRNA. Additionally, the majority of the genes affected in the cell line expressing a lower amount of *PRL* mRNA were included in the list of differentially expressed genes in the first cell line.

This observation suggests that the truncated *PRL* mRNAs could titrate out other regulatory molecules, and hence, the higher the concentration of *PRL* mRNAs, the stronger the effect. One very attractive hypothesis could be that *PRL* mRNAs act as endogenous competing RNAs (ceRNA) and titrate out specific microRNAs [for review see de Giorgio et al. (5)]. However, important investigations are required to unravel the exact molecular mechanism underlying the oncogenic activity of truncated *PRL* mRNAs.

GSEA also indicated a tight link between the genes repressed following *PRL* knockdown in lung cancer cell lines cultured under normoxic conditions and genes that become expressed following hypoxia or a pharmacological stabilization of HIF1 α by a 2-oxoglutarate-dependent dioxygenase inhibitor (7). It has also been shown that HIF1 α contributes to important gene expression programs affecting processes, such as DNA repair, and that a hypoxic microenvironment could enhance genetic alterations in apoptosis-deficient cells (39). By regulating such a specific hypoxia-dependent gene expression program, *PRL* ectopic activation, could confer to cells under normoxic conditions, specific oncogenic properties that normally appear only under hypoxia.

FIG. 5. Identification of *PRL*-responsive genes. (A) In two different *PRL*+ SCLC cell lines (H146 and H524), the transcriptomes of cells treated with control and anti-*PRL* siRNAs were compared to identify upregulated or downregulated genes following *PRL* silencing. The value distribution and mean fold change were calculated and the probe sets corresponding to mRNAs consistently differentially expressed in at least four samples out of the five biological replicates, with a fold change greater than 1.2 retained. After collapsing the probe sets to the genes, a number of differentially expressed genes in both cell lines were represented as Venn diagrams. (B) Validation by RT-qPCR of the 12 differentially expressed genes (common to both cell lines) chosen among those exhibiting a marked fold change after *PRL* knockdown. One hundred percent corresponds to the expression ratio of the values siPRL/siControl obtained after RT-qPCR and normalization with *Actin* and *U6*. Fold changes observed in transcriptomic data are indicated below each histogram. (C) A GSEA was performed on differentially expressed genes according to our transcriptomic approach between H146 SCLC cells expressing *PRL* and the same cells silenced for *PRL* expression (siPRL). In addition to the results shown here, the analysis performed on the C5 group of gene sets, corresponding to genes grouped by gene ontology terms, shows that i/genes downregulated by siPRL (upregulated by *PRL* expression) are significantly enriched in the terms related to actin-based processes, cytoskeleton organization and biogenesis, vesicle and Golgi apparatus-mediated transport, amine and carbohydrate metabolic processes, and oxidoreductase activity; ii/genes upregulated by siPRL (downregulated by *PRL* expression) are enriched in terms related to RNA processing and ribosomal constituents, nuclear factors, mitochondrion, DNA repair, DNA replication, and cell cycle. (D) The inhibition of HDACs by TSA or by SAHA induces a downregulation of *PRL* mRNA. The SCLC cell lines were treated with 100 ng/ml TSA (left panel) or SAHA (right panel) for 6 h. The expression values of the *PRL* gene were obtained from six independent RT-qPCR experiments. The expression values were normalized to the *Actin* control gene and 100% corresponds to the *PRL* expression value in nontreated H146 cells. The significance of the differences in *PRL* expression between treated and untreated conditions was assessed by a Student's *t*-test of mean comparison (***p*-value < 0.05). GSEA, Gene Set Enrichment Analysis; HDAC, histone deacetylase; TSA, trichostatin A. To see this illustration in color, the reader is referred to the web version of this article at www.liebertpub.com/ars



GSEA also revealed a clear and significant overlap between genes upregulated after the knockdown of *PRL* and genes upregulated after a HDAC inhibitor (TSA) treatment or after a targeted knockdown of *HDAC3*, *HDAC2*, and *HDAC1* (18, 33). This observation is reminiscent of the results reported by Schwartz *et al.*, where the knockdown of *BRD4-NUT* in NMC led to a gene activation that was comparable to the effect observed after treatment of cells with HDAC inhibitors (32). In the case of NMC, the molecular dissection of the underlying mechanism showed that the fusion protein *BRD4-NUT* traps CBP/p300 in nuclear foci that can be released after *BRD4-NUT* knockdown (29). Although we have not here been able to identify the molecular basis of the control of gene expression by *PRL*, it is possible to speculate on a similar direct or indirect effect of *PRL* on the activity of major cellular HATs, so that *PRL* knockdown could restore normal acetylation and the underlying gene expression. Additionally, we showed that a TSA treatment can directly downregulate *PRL*. This could also account for the overlap between HDACi gene responses and our cell line response after *PRL* knockdown. The fact that *PRL* is a strong and reliable predictor of poor clinical outcome in the case of lung cancer and that HDAC inhibitors could downregulate *PRL* or activate a shift of the *PRL*-associated aggressive gene signature toward that observed in tumors with better prognosis, suggests that the treatment of *PRL*+ lung tumors with HDACi alone or in combination with a more conventional chemotherapy, should be a reasonable therapeutic strategy for this category of *PRL*+ lung tumor patients.

Finally, this work highlights once more that cancer cells use many yet unknown and unexpected molecular mechanisms to ensure their malignant states and that a comprehensive understanding of cancer cannot be approached without a global consideration of all the oncogenic events.

Materials and Methods

Material

H146, H524, H69, H526, and H1299 cell lines were obtained from the American Type Culture Collection (ATCC). H146, H524, H69, and H526 are derived from small cell lung carcinomas from Caucasian males, respectively, 59, 63, 55, and 55 years old. H1299 is a non-SCLC from a Caucasian 43-year-old male.

Human tissue RNAs were purchased from the Clontech society (Human Total RNA, Master Panel II), except for

placenta RNAs, which were a kind gift from Dr Nadia Al-faidy-Benharouga (CEA).

The Grenoble lung cancer cohort, established by Pr. Elisabeth Brambilla and Christian Brambilla (CHU), includes 293 patients who underwent surgery at the University Hospital of Grenoble. For over 10 years, clinical data (patient age and gender, smoking status, date of diagnosis, histological and TNM classifications of the tumor, treatment received, and date of death) and biological samples (biopsies, tumor sections, and RNAs) were collected. For each patient, a transcriptomic study was performed on a tumor sample collected before implementation of treatment, and in some cases, on a sample of adjacent noncancerous lung tissue. Detailed clinical and biological information relative to these patients is available in Rousseaux *et al.* (30).

PRL+ versus *PRL*− cell line transcriptomes

All transcriptomic data in lung cancer cell lines were obtained using the Illumina HumanHT-12V4.0 expression beadchip and were processed exactly as described in Refs. (14, 28). The raw and processed data from the array are deposited in the GEO databank under the accession GSE49544.

Transcriptomic and gene copy number variants from Grenoble lung cancer tumor cohort

Transcriptomic data from 293 pulmonary tumors of the lung cancer cohort were obtained on the Affymetrix Human Genome U133 Plus 2.0 Array (HG-U133_Plus_2) as described earlier (30). These data are available on the GEO website (reference number GSE30219).

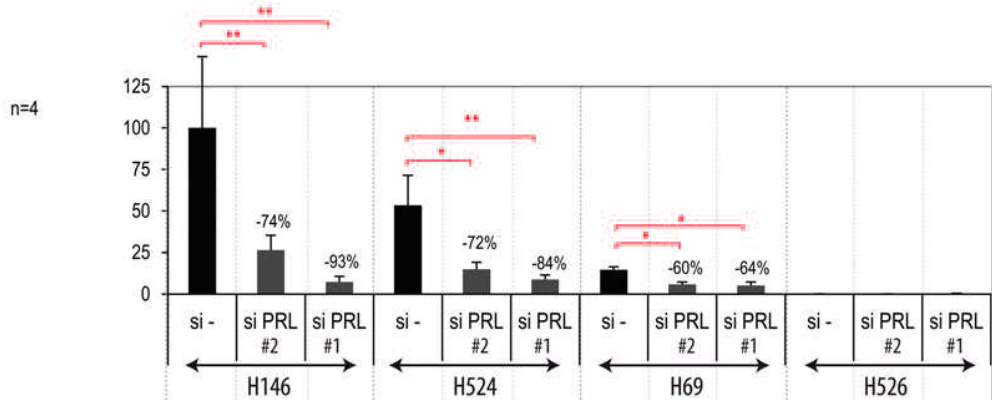
For 247 of the same tumor samples, gene copy numbers were also calculated thanks to parallel hybridization of genomic DNAs on Illumina SNP HumanCNV370 chips (Illumina). The GISTIC version 2.0 algorithm (www.broadinstitute.org/cancer/pub/GISTIC2) was used to identify significant regions of amplification or deletion. The frequencies of aberrations contributing to significant peak regions were compared using chi-square tests.

Analysis of *PRL* and *PRL-R* expression in tumors

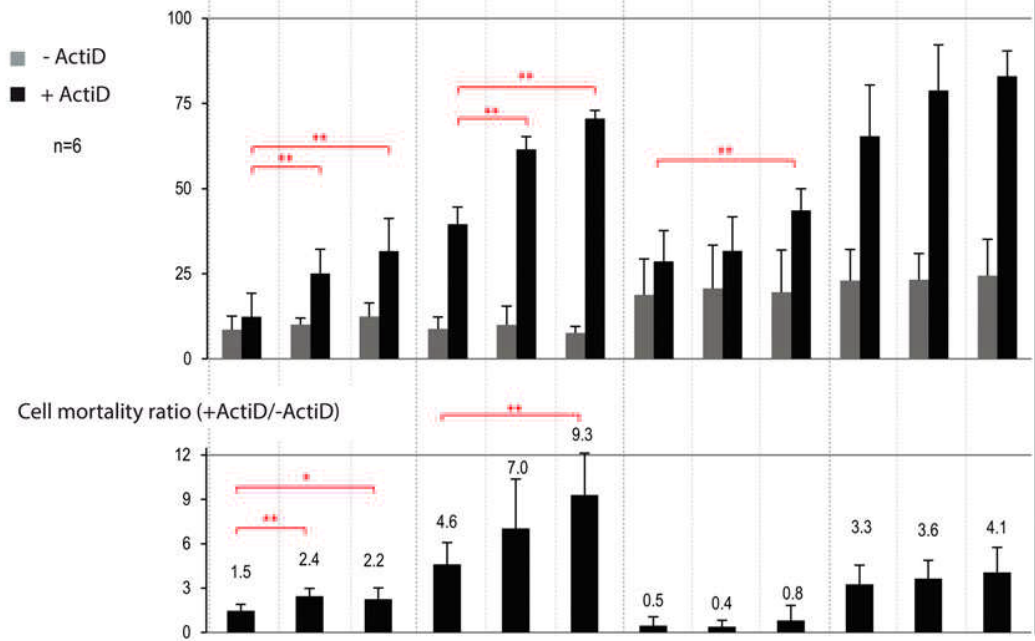
The expression data of *PRL* and *PRL-R* genes in lung tumors were extracted from the Affymetrix RMA normalized signal values corresponding to probe sets 205455_at (*PRL*) and 205455_at (*PRL-R*). For each gene, a threshold value was defined corresponding to a positive expression, by using data

FIG. 6. Effect of *PRL* silencing on cell response to genotoxic treatment in *PRL*+ and *PRL*− SCLC cell lines. SCLC cell lines were transfected with siRNA scrambled (si-) or targeting *PRL* (si*PRL*, two different siRNA duplex) for 48 h, and treated for 12 h by 250 mM Acti D or control treatment. (A) Silencing of the *PRL* gene assessed by RT-qPCR. Expression values of the *PRL* gene were obtained from four independent RT-qPCR experiments. Expressions were normalized to the *Actin* gene. One hundred percent indicates *PRL* expression value in H146 treated with si-Control (scrambled). Silencing efficiencies are indicated. The significance of the differences in *PRL* expression between si- and si*PRL* conditions was assessed by a Student's *t*-test of mean comparison (***p*-value < 0.05; **p*-value < 0.10). (B) (Top panel) The percentage of apoptotic cells was measured by flow cytometry after incubation with an antiactivated caspase-3 antibody in six independent experiments. (Bottom panel). The ratio of cell mortality between Acti D-treated and nontreated cells was calculated for each of the six experiments. The significance of the differences in the apoptotic cell proportions between si- and si*PRL* conditions was assessed by a Student's *t*-test of mean comparison (***p*-value < 0.05; **p*-value < 0.10). (C) Example of flow cytometry data. Each dot represents one H524 cell. The PE fluorescence of the antiactivated caspase-3 staining is represented on the y-axis. The FSC size parameter is represented on the x-axis. The P3 delimited region comprises apoptotic cells. Acti D, actinomycin D; FSC, forward-scattered light; PE, phycoerythrin. To see this illustration in color, the reader is referred to the web version of this article at www.liebertpub.com/ars

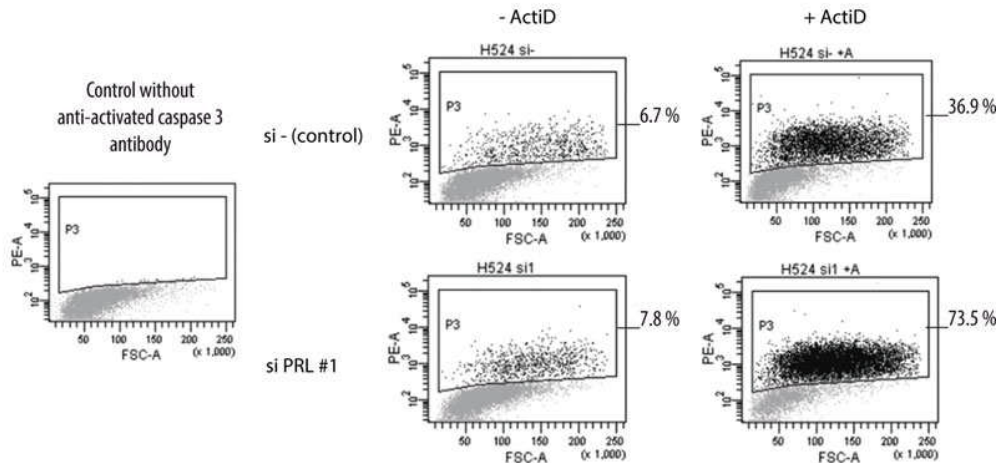
A PRL expression ratio



B Cell mortality (% activated caspase 3 positive cells)



C



from a series of normal human somatic tissue samples, as described previously (30). For the *PRL* gene, the threshold above which it was considered expressed was the mean value of the *205455_at* probe set signal in 109 samples of normal human somatic tissues (excluding the pituitary samples). For the *PRL-R* gene, the threshold above which it was considered expressed was the mean + 2 standard deviations of the *205455_at* of the signal values in 112 normal human somatic tissue samples.

RNA extraction and reverse transcription

SCLC cells growing in suspension were harvested by pipetting 2 days after seeding (and eventual transfection of siRNA or drug treatment). The extraction of total RNA was performed with the NucleoSpin RNA II (Macherey-Nagel) kit according to the supplier's recommendations. The RNA content and purity were assessed using NanoDrop 1000 (Thermo Scientific). The reverse transcription reaction was performed on 1 μ g of total RNA using the kit SuperScript III RT First-Strand Synthesis SuperMix for qRT-PCR (Invitrogen) according to the supplier's recommendations.

PCR reactions

Primers were designed with the help of the three following online software: Universal Probe Library Assay Design Center (Roche Applied Science), Primer 3 (<http://frodo.wi.mit.edu/>), or Primer Bank (<http://pga.mgh.harvard.edu/primerbank/>), and their specificity was checked using BLAST (Basic Local Alignment Search Tool) (<http://blast.ncbi.nlm.nih.gov/Blast.cgi>). The used primers and their sequences are listed in Supplementary Table S1B.

The PCR reactions were carried out in a final volume of 25 μ l containing 0.5 μ l dNTPs (1 mM, 5'), 0.5 μ l of each primer, forward and reverse (10 μ M), 2.5 μ l of 10 \times buffer, 0.5 μ l of TaqPol enzyme (5 U/ μ l, 5'), and cDNAs (RT product diluted 1/10). The PCR program was as follows: denaturation at 95°C for 1 min, then 35 cycles (denaturation 95°C 30 s, annealing 30 s, elongation 68°C), and finally 5 min of terminal extension at 68°C. The elongation time varied depending on the length of the PCR product expected (1 min/kb). The annealing temperature was chosen according to the T_m of primers (often 65°C, sometimes down to 55°C). A negative control was achieved by replacing the cDNA by DEPC H₂O. The PCR products were analyzed by electrophoresis in a gel containing 1% agarose (Invitrogen) and 5 μ l of BET (10 mg/ml; Sigma-Aldrich) in 100 ml of 1 \times TAE buffer. DNA (25 μ l of PCR product) was mixed with 3 μ l of loading buffer (bromophenol blue, 40% sucrose) and loaded onto the gel, together with a size marker (DNA Ladder, Fermentas or 1 kb-Plus; Invitrogen). Migration was performed in 1 \times TAE at 100 V for about 30 min. The reading was done under a UV lamp.

qPCR reactions

qPCR analysis was performed using Minimum Information for Publication of Quantitative Real-Time PCR Experiments (MIQE) guidelines (3). Primers were designed so that the length of the amplicon is between 70 and 200 bp. qPCR was performed in 96-well plates in a final volume of 20 μ l. Each well contained 0.03 μ l of each primer, forward and

reverse (100 μ M), 10 μ l of Brilliant SYBR Green 2 \times III (Agilent Technology), and 4 μ l of cDNA (RT product diluted 1/10). For each condition, a technical duplicate and a control point where the cDNA is replaced by the extracted RNA (no RT) were added. The reaction was carried out in the Mx3005P qPCR cycloer (Agilent Technology) with the following program: 3 min of initial denaturation at 95°C, then 40 cycles with 20 s of denaturation at 95°C, and 20 s of hybridization and elongation at 60°C.

The qPCR analysis was performed using the MxPro software (Agilent Technology). After verifying that a given pair of primers gives a single peak of dissociation, the value of the threshold cycle (C_t) was given by the number of cycles necessary for the measured fluorescence value to exceed the threshold value (threshold = 235). When the C_t exceeded 35, we considered that there was no expression. When the C_t value obtained for the no RT control point was too close (NoRT-RT < 2), we did not take the value into account. The analysis was performed on the average of the two C_t values obtained for the two technical replicates (difference < 1). Expression values were normalized with the selected control genes (*Actin*, *U6*, *RELA*, and/or *AUP1*). The expression values were given as a percentage of the value corresponding to a reference condition (reference tissue such as pituitary or placenta or breast, or nontransfected control condition for cell lines) by the method of Delta C_t (18): $2^{\Delta C_t}$ gene of interest (reference condition – condition of interest) / $2^{\Delta C_t}$ control gene (reference condition – condition of interest).

Cell culture and transfection

H524 and H526 cells were cultured in the Roswell Park Memorial Institute medium (RPMI 1640; Gibco) supplemented with 10% fetal bovine serum (PAA), 2 mM L-glutamine (Gibco), and 1% penicillin–streptomycin (Gibco). For H69 cells, the complete RPMI medium was supplemented with 10 mM HEPES (Gibco), 1 mM sodium pyruvate (Gibco), and 2.5 g/L glucose (Gibco). For H146 cells, the complete RPMI medium was supplemented with 10 mM HEPES (Gibco), 1 mM sodium pyruvate (Gibco), 4.5 g/L glucose (Gibco), and 1.5 g/L sodium bicarbonate (Gibco). Cells were maintained in 75-cm² flasks at 37°C in a humidified 5% CO₂ incubator.

siRNAs targeting human *PRL* and nonspecific control siRNAs were designed with the DSIR software (<http://biodev.extra.cea.fr/DSIR/>). *PRL* siRNA sequences are shown in Supplementary Table S1A. Transfections were carried out in six-well plates with 1.5 μ l siRNA duplex (20 μ M), 5 μ l Lipofectamine RNAiMAX Reagent (Invitrogen), 500 μ l of Opti-MEM medium (Gibco), and 250,000 cells in an antibiotic-free culture medium.

Two days after transfection, the culture medium was complemented with 250 nM Acti D (Sigma-Aldrich) for 12 h or 100 ng/ml TSA (Sigma-Aldrich) for 6 h.

Quantification of apoptotic cells

Cells were harvested and washed in PBS. Active caspase-3 was detected using the Active caspase-3 antibody kit (BD Pharmingen) following the manufacturer's instructions. Briefly, cells were fixed and permeabilized for 30 min, and then incubated with the phycoerythrin-conjugated antiactive caspase-3 antibody. Analysis of active caspase-3-positive

cells was performed by counting 10,000 events on a Becton Dickinson FACScan flow cytometer and data were fit using DIVA6 software (Becton Dickinson).

Acknowledgments

The authors wish to thank Dr Nadia Alfaidy-Benharouga (CEA, Grenoble) for her kind gift of human placenta RNA samples. The work in S.K. and E.B. laboratories is supported by INCa, ANR EpiSperm, and ARC Subvention libre programs. A.D. has been fully supported by INCa, AGIRDOM, and ANR grants. The clinical research on lung cancer was funded by PNES POU MON INCA 2005 and BIOMARK-SCAN PHRC 2003. The transcriptomic analyses were performed thanks to the program Carte d'Identité des Tumeurs supported by the Ligue NATIONALE CONTRE LE CANCER.

Author Disclosure Statement

The authors declare that there is no competing financial interest.

References

1. Axelsen JB, Lotem J, Sachs L, and Domany E. Genes overexpressed in different human solid cancers exhibit different tissue-specific expression profiles. *Proc Natl Acad Sci U S A* 104: 13122–13127, 2007.
2. Bole-Feysot C, Goffin V, Edery M, Binart N, and Kelly PA. Prolactin (PRL) and its receptor: actions, signal transduction pathways and phenotypes observed in PRL receptor knockout mice. *Endocr Rev* 19: 225–268, 1998.
3. Bustin SA, Benes V, Garson JA, Hellemans J, Huggett J, Kubista M, Mueller R, Nolan T, Pfaffl MW, Shipley GL, Vandesompele J, and Wittwer CT. The MIQE guidelines: minimum information for publication of quantitative real-time PCR experiments. *Clin Chem* 55: 611–622, 2009.
4. Chen YT, Scanlan MJ, Venditti CA, Chua R, Theiler G, Stevenson BJ, Iseli C, Gure AO, Vasicek T, Strausberg RL, Jongeneel CV, Old LJ, and Simpson AJ. Identification of cancer/testis-antigen genes by massively parallel signature sequencing. *Proc Natl Acad Sci U S A* 102: 7940–7945, 2005.
5. de Giorgio A, Krell J, Harding V, Stebbing J, and Castellano L. Emerging roles of competing endogenous RNAs in cancer: insights from the regulation of PTEN. *Mol Cell Biol* 33: 3976–3982, 2013.
6. Doyle JM, Gao J, Wang J, Yang M, and Potts PR. MAGE-RING protein complexes comprise a family of E3 ubiquitin ligases. *Mol Cell* 39: 963–974, 2010.
7. Elvidge GP, Glenny L, Appelhoff RJ, Ratcliffe PJ, Ragoussis J, and Gleadle JM. Concordant regulation of gene expression by hypoxia and 2-oxoglutarate-dependent dioxygenase inhibition: the role of HIF-1 α , HIF-2 α , and other pathways. *J Biol Chem* 281: 15215–15226, 2006.
8. Emadali A, Rousseaux S, Bruder-Costa J, Rome C, Duley S, Hamaidia S, Betton P, Debernardi A, Leroux D, Bernay B, Kieffer-Jaquinod S, Combes F, Ferri E, McKenna CE, Petosa C, Bruley C, Garin J, Ferro M, Gressin R, Callanan MB, and Khochbin S. Identification of a novel BET bromodomain inhibitor-sensitive, gene regulatory circuit that controls Rituximab response and tumour growth in aggressive lymphoid cancers. *EMBO Mol Med* 5: 1180–1195, 2013.
9. Feichtinger J, Aldeaij I, Anderson R, Almutairi M, Almatrafi A, Alsiwiehri N, Griffiths K, Stuart N, Wakeman JA, Larcombe L, and McFarlane RJ. Meta-analysis of clinical data using human meiotic genes identifies a novel cohort of highly restricted cancer-specific marker genes. *Oncotarget* 3: 843–853, 2012.
10. Fernandez I, Touraine P, and Goffin V. Prolactin and human tumorigenesis. *J Neuroendocrinol* 22: 771–777, 2010.
11. Filippakopoulos P, Qi J, Picaud S, Shen Y, Smith WB, Fedorov O, Morse EM, Keates T, Hickman TT, Felletar I, Philpott M, Munro S, McKeown MR, Wang Y, Christie AL, West N, Cameron MJ, Schwartz B, Heightman TD, La Thangue N, French CA, Wiest O, Kung AL, Knapp S, and Bradner JE. Selective inhibition of BET bromodomains. *Nature* 468: 1067–1073, 2010.
12. French CA. Pathogenesis of NUT midline carcinoma. *Annu Rev Pathol* 7: 247–265, 2012.
13. Gandhi L and Johnson BE. Paraneoplastic syndromes associated with small cell lung cancer. *J Natl Compr Canc Netw* 4: 631–638, 2006.
14. Gaucher J, Boussouar F, Montellier E, Curtet S, Buchou T, Bertrand S, Hery P, Jounier S, Depaux A, Vitte AL, Guardiola P, Pernet K, Debernardi A, Lopez F, Holota H, Imbert J, Wolgemuth DJ, Gerard M, Rousseaux S, and Khochbin S. Bromodomain-dependent stage-specific male genome programming by Brdt. *EMBO J* 31: 3809–3820, 2012.
15. Gerlo S, Davis JR, Mager DL, and Kooijman R. Prolactin in man: a tale of two promoters. *Bioessays* 28: 1051–1055, 2006.
16. Goffin V and Carter-Su C. Meeting report: the 2012 FASEB Science Research Conference “The growth hormone/prolactin family in biology and disease” A novel biannual rendez-vous in the endocrinology landscape. *Pediatr Endocrinol Rev* 10: 243–245, 2012.
17. Hanahan D and Weinberg RA. Hallmarks of cancer: the next generation. *Cell* 144: 646–674, 2013.
18. Heller G, Schmidt WM, Ziegler B, Holzer S, Mullauer L, Bilban N, Zielinski CC, Drach J, and Zochbauer-Muller S. Genome-wide transcriptional response to 5-aza-2'-deoxycytidine and trichostatin A in multiple myeloma cells. *Cancer Res* 68: 44–54, 2008.
19. Hofmann O, Caballero OL, Stevenson BJ, Chen YT, Cohen T, Chua R, Maher CA, Panji S, Schaefer U, Kruger A, Lehvaslaiho M, Carninci P, Hayashizaki Y, Jongeneel CV, Simpson AJ, Old LJ, and Hide W. Genome-wide analysis of cancer/testis gene expression. *Proc Natl Acad Sci U S A* 105: 20422–20427, 2008.
20. Ignacak A, Kasztelnik M, Sliwa T, Korbut RA, Rajda K, and Guzik TJ. Prolactin—not only lactotrophin. A “new” view of the “old” hormone. *J Physiol Pharmacol* 63: 435–443, 2012.
21. Janic A, Mendizabal L, Llamazares S, Rossell D, and Gonzalez C. Ectopic expression of germline genes drives malignant brain tumor growth in *Drosophila*. *Science* 330: 1824–1827, 2011.
22. Jazirehi AR. Regulation of apoptosis-associated genes by histone deacetylase inhibitors: implications in cancer therapy. *Anticancer Drugs* 21: 805–813, 2010.
23. Lotem J, Netanel D, Domany E, and Sachs L. Human cancers overexpress genes that are specific to a variety of

- normal human tissues. *Proc Natl Acad Sci U S A* 102: 18556–18561, 2005.
24. Lotem J and Sachs L. Epigenetics and the plasticity of differentiation in normal and cancer stem cells. *Oncogene* 25: 7663–7672, 2006.
 25. Mathieu R, Evrard B, Fromont G, Rioux-Leclercq N, Godet J, Cathelineau X, Guille F, Primig M, and Chalmel F. Expression screening of cancer/testis genes in prostate cancer identifies NR6A1 as a novel marker of disease progression and aggressiveness. *Prostate* 73: 1103–1114, 2013.
 26. Mehnert JM and Kelly WK. Histone deacetylase inhibitors: biology and mechanism of action. *Cancer J* 13: 23–29, 2007.
 27. Monte M, Simonatto M, Peche LY, Bublik DR, Gobessi S, Pierotti MA, Rodolfo M, and Schneider C. MAGE-A tumor antigens target p53 transactivation function through histone deacetylase recruitment and confer resistance to chemotherapeutic agents. *Proc Natl Acad Sci U S A* 103: 11160–11165, 2006.
 28. Montellier E, Boussouar F, Rousseaux S, Zhang K, Buchou T, Fenaille F, Shiota H, Debernardi A, Hery P, Curtet S, Jamshidikia M, Barral S, Holota H, Bergon A, Lopez F, Guardiola P, Pernet K, Imbert J, Petosa C, Tan M, Zhao Y, Gerard M, and Khochbin S. Chromatin-to-nucleoprotamine transition is controlled by the histone H2B variant TH2B. *Genes Dev* 27: 1680–1692, 2013.
 29. Reynoird N, Schwartz BE, Delvecchio M, Sadoul K, Meyers D, Mukherjee C, Caron C, Kimura H, Rousseaux S, Cole PA, Panne D, French CA, and Khochbin S. Oncogenesis by sequestration of CBP/p300 in transcriptionally inactive hyperacetylated chromatin domains. *EMBO J* 29: 2943–2952, 2010.
 30. Rousseaux S, Debernardi A, Jacquiau B, Vitte AL, Vesin A, Nagy-Mignotte H, Moro-Sibilot D, Brichon PY, Lantuejoul S, Hainaut P, Laffaire J, de Reynies A, Beer DG, Timsit JF, Brambilla C, Brambilla E, and Khochbin S. Ectopic activation of germline and placental genes identifies aggressive metastasis-prone lung cancers. *Sci Transl Med* 5: 186ra66, 2013.
 31. Rousseaux S, Wang J, and Khochbin S. Cancer hallmarks sustained by ectopic activations of placenta/male germline genes. *Cell Cycle* 12: 2331–2332, 2013.
 32. Schwartz BE, Hofer MD, Lemieux ME, Bauer DE, Cameron MJ, West NH, Agoston ES, Reynoird N, Khochbin S, Ince TA, Christie A, Janeway KA, Vargas SO, Perez-Atayde AR, Aster JC, Sallan SE, Kung AL, Bradner JE, and French CA. Differentiation of NUT midline carcinoma by epigenomic reprogramming. *Cancer Res* 71: 2686–2696, 2011.
 33. Senese S, Zaragoza K, Minardi S, Muradore I, Ronzoni S, Passafaro A, Bernard L, Draetta GF, Alcalay M, Seiser C, and Chiocca S. Role for histone deacetylase 1 in human tumor cell proliferation. *Mol Cell Biol* 27: 4784–4795, 2007.
 34. Shen H and Laird PW. Interplay between the cancer genome and epigenome. *Cell* 153: 38–55, 2013.
 35. van Meerbeeck JP, Fennell DA, and De Ruysscher DK. Small-cell lung cancer. *Lancet* 378: 1741–1755, 2011.
 36. Wang J, Emadali A, Le Bescont A, Callanan M, Rousseaux S, and Khochbin S. Induced malignant genome reprogramming in somatic cells by testis-specific factors. *Biochim Biophys Acta* 1809: 221–225, 2011.
 37. Wang J, Rousseaux S, and Khochbin S. Sustaining cancer through additive ectopic gene activation. *Curr Opin Oncol* 26: 73–77, 2013.
 38. Yang B, O'Herrin S, Wu J, Reagan-Shaw S, Ma Y, Nihal M, and Longley BJ. Select cancer testis antigens of the MAGE-A, -B, and -C families are expressed in mast cell lines and promote cell viability *in vitro* and *in vivo*. *J Invest Dermatol* 127: 267–275, 2007.
 39. Yoo YG, Hayashi M, Christensen J, and Huang LE. An essential role of the HIF-1 α -c-Myc axis in malignant progression. *Ann N Y Acad Sci* 1177: 198–204, 2009.

Address correspondence to:

Dr. Sophie Rousseaux

INSERM U823; Université Joseph Fourier–Grenoble 1

Institut Albert Bonniot

Grenoble

France

E-mail: sophie.rousseau@ujf-grenoble.fr

Dr. Saadi Khochbin

INSERM U823; Université Joseph Fourier–Grenoble 1

Institut Albert Bonniot

Grenoble

France

E-mail: khochbin@ujf-grenoble.fr

Date of first submission to ARS Central, August 8, 2013; date of final revised submission, February 3, 2014; date of acceptance, February 8, 2014.

Abbreviations Used

Acti D = actinomycin D
 CGH = comparative genomic hybridization
 DLBCL = diffuse large B-cell lymphoma
 EST = expressed sequence tags
 FSC = forward-scattered light
 GSEA = Gene Set Enrichment Analysis
 HDAC = histone deacetylase
 HDACi = histone deacetylase inhibitor
 LCNE = large cell neuroendocrine
 NMC = NUT midline carcinoma
 PE = phycoerythrin
 PRL = prolactin
 PRL-R = prolactin receptor
 (q)PCR = (quantitative) polymerase chain reaction
 RT = reverse transcription
 SCLC = small cell lung cancer
 TNM = tumor node metastasis
 TSA = trichostatin A
 UTR = untranslated region

BIBLIOGRAPHIE

- 96-1133, D.n. 1996. Décret n° 96-1133 du 24 décembre 1996 relatif à l'interdiction de l'amiante, pris en application du code du travail et du code de la consommation Vol. Décret n° 96-1133 du 24 décembre 1996.
- Abedin, Z., M. Louis-Juste, M. Stangl, and J. Field. 2013. The role of base excision repair genes OGG1, APN1 and APN2 in benzo[a]pyrene-7,8-dione induced p53 mutagenesis. *Mutation research*. 750:121-128.
- Adler, I. 1912. Primary Malignant Growths of the Lungs and Bronchi.
- Allison, M. 2010. Turning the tide in lung cancer. *Nature biotechnology*. 28:999-1002.
- Almaric, F. 2007. Analyse économique des coûts du cancer. Institut national du cancer.
- An, Q., Y. Liu, Y. Gao, J. Huang, X. Fong, L. Liu, D. Zhang, J. Zhang, and S. Cheng. 2002. Deletion of tumor suppressor genes in Chinese non-small cell lung cancer. *Cancer letters*. 184:189-195.
- Anderson, E., J.E. Ferguson, H. Morten, S.M. Shalet, E.L. Robinson, and A. Howell. 1993. Serum immunoreactive and bioactive lactogenic hormones in advanced breast cancer patients treated with bromocriptine and octreotide. *European journal of cancer*. 29A:209-217.
- Aprelikova, O., S. Pandolfi, S. Tackett, M. Ferreira, K. Salnikow, Y. Ward, J.I. Risinger, J.C. Barrett, and J. Niederhuber. 2009. Melanoma antigen-11 inhibits the hypoxia-inducible factor prolyl hydroxylase 2 and activates hypoxic response. *Cancer research*. 69:616-624.
- Arendt, L.M., D.E. Rugowski, T.A. Grafwallner-Huseth, M.J. Garcia-Barchino, H. Rui, and L.A. Schuler. 2011. Prolactin-induced mouse mammary carcinomas model estrogen resistant luminal breast cancer. *Breast cancer research : BCR*. 13:R11.
- Arents, G., R.W. Burlingame, B.C. Wang, W.E. Love, and E.N. Moudrianakis. 1991. The nucleosomal core histone octamer at 3.1 Å resolution: a tripartite protein assembly and a left-handed superhelix. *Proceedings of the National Academy of Sciences of the United States of America*. 88:10148-10152.
- Arias, J., A.S. Alberts, P. Brindle, F.X. Claret, T. Smeal, M. Karin, J. Feramisco, and M. Montminy. 1994. Activation of cAMP and mitogen responsive genes relies on a common nuclear factor. *Nature*. 370:226-229.
- Arif, M., P. Senapati, J. Shandilya, and T.K. Kundu. 2010. Protein lysine acetylation in cellular function and its role in cancer manifestation. *Biochimica et biophysica acta*. 1799:702-716.
- Ariyoshi, N., M. Miyamoto, Y. Umetsu, H. Kunitoh, H. Dosaka-Akita, Y. Sawamura, J. Yokota, N. Nemoto, K. Sato, and T. Kamataki. 2002. Genetic polymorphism of CYP2A6 gene and tobacco-induced lung cancer risk in male smokers. *Cancer epidemiology, biomarkers & prevention : a publication of the American Association for Cancer Research, cosponsored by the American Society of Preventive Oncology*. 11:890-894.
- Artandi, S.E., and R.A. DePinho. 2000. A critical role for telomeres in suppressing and facilitating carcinogenesis. *Current opinion in genetics & development*. 10:39-46.
- Auerbach, O., A.P. Stout, E.C. Hammond, and L. Garfinkel. 1961. Changes in bronchial epithelium in relation to cigarette smoking and in relation to lung cancer. *The New England journal of medicine*. 265:253-267.
- Axelsen, J.B., J. Lotem, L. Sachs, and E. Domany. 2007. Genes overexpressed in different human solid cancers exhibit different tissue-specific expression profiles. *Proceedings of the National Academy of Sciences of the United States of America*. 104:13122-13127.
- Ayyanathan, K., M.S. Lechner, P. Bell, G.G. Maul, D.C. Schultz, Y. Yamada, K. Tanaka, K. Torigoe, and F.J. Rauscher, 3rd. 2003. Regulated recruitment of HP1 to a euchromatic gene induces mitotically heritable, epigenetic gene silencing: a mammalian cell culture model of gene variegation. *Genes & development*. 17:1855-1869.
- Azim, H.A., Jr., and A.K. Ganti. 2007. Treatment options for relapsed small-cell lung cancer. *Anti-cancer drugs*. 18:255-261.
- Bachman, K.E., M.R. Rountree, and S.B. Baylin. 2001. Dnmt3a and Dnmt3b are transcriptional repressors that exhibit unique localization properties to heterochromatin. *The Journal of biological chemistry*. 276:32282-32287.
- Bailey-Wilson, J.E., C.I. Amos, S.M. Pinney, G.M. Petersen, M. de Andrade, J.S. Wiest, P. Fain, A.G. Schwartz, M. You, W. Franklin, C. Klein, A. Gazdar, H. Rothschild, D. Mandal, T. Coons, J. Slusser, J. Lee, C. Gaba, E. Kupert, A. Perez, X. Zhou, D. Zeng, Q. Liu, Q. Zhang, D. Seminara, J. Minna, and M.W. Anderson. 2004. A major lung cancer susceptibility locus maps to chromosome 6q23-25. *American journal of human genetics*. 75:460-474.
- Bannister, A.J., and T. Kouzarides. 1996. The CBP co-activator is a histone acetyltransferase. *Nature*. 384:641-643.
- Barlesi, F., G. Giaccone, M.I. Gallegos-Ruiz, A. Loundou, S.W. Span, P. Lefesvre, F.A. Krut, and J.A. Rodriguez. 2007. Global histone modifications predict prognosis of resected non small-cell lung cancer. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology*. 25:4358-4364.
- Barski, A., S. Cuddapah, K. Cui, T.Y. Roh, D.E. Schones, Z. Wang, G. Wei, I. Chepelev, and K. Zhao. 2007. High-resolution profiling of histone methylations in the human genome. *Cell*. 129:823-837.
- Bartling, B., H.S. Hofmann, T. Boettger, G. Hansen, S. Burdach, R.E. Silber, and A. Simm. 2005. Comparative application of antibody and gene array for expression profiling in human squamous cell lung carcinoma. *Lung cancer*. 49:145-154.
- Baudat, F., and B. de Massy. 2004. [SPO11: an activity that promotes DNA breaks required for meiosis]. *Medecine sciences : M/S*. 20:213-218.
- Baylin, S.B., J.W. Hoppener, A. de Bustros, P.H. Steenbergh, C.J. Lips, and B.D. Nelkin. 1986. DNA methylation patterns of the calcitonin gene in human lung cancers and lymphomas. *Cancer research*. 46:2917-2922.

- Bazan, J.F. 1990. Structural design and molecular evolution of a cytokine receptor superfamily. *Proceedings of the National Academy of Sciences of the United States of America*. 87:6934-6938.
- Beato, M., and K. Eisfeld. 1997. Transcription factor access to chromatin. *Nucleic acids research*. 25:3559-3563.
- Belinsky, S.A., D.M. Klinge, J.D. Dekker, M.W. Smith, T.J. Bocklage, F.D. Gilliland, R.E. Crowell, D.D. Karp, C.A. Stidley, and M.A. Picchi. 2005. Gene promoter methylation in plasma and sputum increases with lung cancer risk. *Clinical cancer research : an official journal of the American Association for Cancer Research*. 11:6505-6511.
- Belinsky, S.A., D.M. Klinge, C.A. Stidley, J.P. Issa, J.G. Herman, T.H. March, and S.B. Baylin. 2003. Inhibition of DNA methylation and histone deacetylation prevents murine lung cancer. *Cancer research*. 63:7089-7093.
- Belinsky, S.A., K.J. Nikula, W.A. Palmisano, R. Michels, G. Saccomanno, E. Gabrielson, S.B. Baylin, and J.G. Herman. 1998. Aberrant methylation of p16(INK4a) is an early event in lung cancer and a potential biomarker for early diagnosis. *Proceedings of the National Academy of Sciences of the United States of America*. 95:11891-11896.
- Belinsky, S.A., W.A. Palmisano, F.D. Gilliland, L.A. Crooks, K.K. Divine, S.A. Winters, M.J. Grimes, H.J. Harms, C.S. Tellez, T.M. Smith, P.P. Moots, J.F. Lechner, C.A. Stidley, and R.E. Crowell. 2002. Aberrant promoter methylation in bronchial epithelium and sputum from current and former smokers. *Cancer research*. 62:2370-2377.
- Belot, A., P. Grosclaude, N. Bossard, E. Jouglu, E. Benhamou, P. Delafosse, A.V. Guizard, F. Molinie, A. Danzon, S. Bara, A.M. Bouvier, B. Tretarre, F. Binder-Foucard, M. Colonna, L. Daubisse, G. Hedelin, G. Launoy, N. Le Stang, M. Maynadie, A. Monnereau, X. Troussard, J. Faivre, A. Collignon, I. Janoray, P. Arveux, A. Buemi, N. Raverdy, C. Schvartz, M. Bovet, L. Cherie-Challine, J. Esteve, L. Remontet, and M. Velten. 2008. Cancer incidence and mortality in France over the period 1980-2005. *Revue d'epidemiologie et de sante publique*. 56:159-175.
- Ben-Ezra, J.M., M.J. Kornstein, M.M. Grimes, and G. Krystal. 1994. Small cell carcinomas of the lung express the Bcl-2 protein. *The American journal of pathology*. 145:1036-1040.
- Ben-Jonathan, N., K. Liby, M. McFarland, and M. Zinger. 2002. Prolactin as an autocrine/paracrine growth factor in human cancer. *Trends in endocrinology and metabolism: TEM*. 13:245-250.
- Ben-Jonathan, N., J.L. Mershon, D.L. Allen, and R.W. Steinmetz. 1996. Extrapituitary prolactin: distribution, regulation, functions, and clinical aspects. *Endocrine reviews*. 17:639-669.
- Bentley, P. 1999. Genetically modified cells, or how some cancers occur. *Nursing times*. 95:48-49.
- Berdasco, M., and M. Esteller. 2010. Aberrant epigenetic landscape in cancer: how cellular identity goes awry. *Developmental cell*. 19:698-711.
- Bernichtein, S., P. Touraine, and V. Goffin. 2010. New concepts in prolactin biology. *The Journal of endocrinology*. 206:1-11.
- Berrino, F., R. De Angelis, M. Sant, S. Rosso, M. Bielska-Lasota, J.W. Coebergh, M. Santaquilani, and E.W. group. 2007. Survival for eight major cancers and all cancers combined for European adults diagnosed in 1995-99: results of the EUROCARE-4 study. *The lancet oncology*. 8:773-783.
- Betticher, D.C. 2005. Adjuvant and neoadjuvant chemotherapy in NSCLC: a paradigm shift. *Lung cancer*. 50 Suppl 2:S9-16.
- Bhan, S., S.S. Negi, C. Shao, C.A. Glazer, A. Chuang, D.A. Gaykalova, W. Sun, D. Sidransky, P.K. Ha, and J.A. Califano. 2011. BORIS binding to the promoters of cancer testis antigens, MAGEA2, MAGEA3, and MAGEA4, is associated with their transcriptional activation in lung cancer. *Clinical cancer research : an official journal of the American Association for Cancer Research*. 17:4267-4276.
- Bhatavdekar, J.M., D.D. Patel, N.H. Karelia, N.G. Shah, N. Ghosh, H.H. Vora, T.P. Suthar, D.B. Balar, and S.S. Doctor. 1994. Can plasma prolactin predict tamoxifen resistance in patients with advanced breast cancer? *European journal of surgical oncology : the journal of the European Society of Surgical Oncology and the British Association of Surgical Oncology*. 20:118-121.
- Bhattacharya, S.K., S. Ramchandani, N. Cervoni, and M. Szyf. 1999. A mammalian protein with specific demethylase activity for mCpG DNA. *Nature*. 397:579-583.
- Bhoite, L.T., Y. Yu, and D.J. Stillman. 2001. The Swi5 activator recruits the Mediator complex to the HO promoter without RNA polymerase II. *Genes & development*. 15:2457-2469.
- Bieche, I., C. Nogues, V. Paradis, M. Olivi, P. Bedossa, R. Lidereau, and M. Vidaud. 2000. Quantitation of hTERT gene expression in sporadic breast tumors with a real-time reverse transcription-polymerase chain reaction assay. *Clinical cancer research : an official journal of the American Association for Cancer Research*. 6:452-459.
- Binart, N., A. Bachelot, and J. Bouilly. 2010. Impact of prolactin receptor isoforms on reproduction. *Trends in endocrinology and metabolism: TEM*. 21:362-368.
- Bird, A. 2002. DNA methylation patterns and epigenetic memory. *Genes & development*. 16:6-21.
- Bird, A.P., and A.P. Wolffe. 1999. Methylation-induced repression--belts, braces, and chromatin. *Cell*. 99:451-454.
- Black, C., R. de Verteuil, S. Walker, J. Ayres, A. Boland, A. Bagust, and N. Waugh. 2007. Population screening for lung cancer using computed tomography, is there evidence of clinical effectiveness? A systematic review of the literature. *Thorax*. 62:131-138.
- Blackwood, E.M., and J.T. Kadonaga. 1998. Going the distance: a current view of enhancer action. *Science*. 281:60-63.
- Bock, C., E.M. Tomazou, A.B. Brinkman, F. Muller, F. Simmer, H. Gu, N. Jager, A. Gnirke, H.G. Stunnenberg, and A. Meissner. 2010. Quantitative comparison of genome-wide DNA methylation mapping technologies. *Nature biotechnology*. 28:1106-1114.
- Bodner, S.M., J.D. Minna, S.M. Jensen, D. D'Amico, D. Carbone, T. Mitsudomi, J. Fedorko, D.L. Buchhagen, M.M. Nau, A.F. Gazdar, and et al. 1992. Expression of mutant p53 proteins in lung cancer correlates with the class of p53 gene mutation. *Oncogene*. 7:743-749.

- Boeger, H., J. Griesenbeck, J.S. Strattan, and R.D. Kornberg. 2003. Nucleosomes unfold completely at a transcriptionally active promoter. *Molecular cell*. 11:1587-1598.
- Boel, P., C. Wildmann, M.L. Sensi, R. Brasseur, J.C. Renauld, P. Coulie, T. Boon, and P. van der Bruggen. 1995. BAGE: a new gene encoding an antigen recognized on human melanomas by cytolytic T lymphocytes. *Immunity*. 2:167-175.
- Bole-Feyssot, C., V. Goffin, M. Edery, N. Binart, and P.A. Kelly. 1998. Prolactin (PRL) and its receptor: actions, signal transduction pathways and phenotypes observed in PRL receptor knockout mice. *Endocrine reviews*. 19:225-268.
- Bonnerterre, J., L. Mauriac, B. Weber, H. Roche, P. Fargeot, M. Tubiana-Hulin, M. Sevin, P. Chollet, and P. Cappelaere. 1988. Tamoxifen plus bromocriptine versus tamoxifen plus placebo in advanced breast cancer: results of a double blind multicentre clinical trial. *European journal of cancer & clinical oncology*. 24:1851-1853.
- Boutin, J.M., C. Jolicoeur, H. Okamura, J. Gagnon, M. Edery, M. Shirota, D. Banville, I. Dusanter-Fourt, J. Djiane, and P.A. Kelly. 1988. Cloning and expression of the rat prolactin receptor, a member of the growth hormone/prolactin receptor gene family. *Cell*. 53:69-77.
- Bowen, N.J., M.B. Palmer, and P.A. Wade. 2004. Chromosomal regulation by MeCP2: structural and enzymatic considerations. *Cellular and molecular life sciences : CMLS*. 61:2163-2167.
- Brambilla, E., S. Gazzeri, D. Moro, C. Caron de Fromental, V. Gouyer, M. Jacrot, and C. Brambilla. 1993. Immunohistochemical study of p53 in human lung carcinomas. *The American journal of pathology*. 143:199-210.
- Brambilla, E., V. Gouyer, S. Gazzieri, I. Bolloni, and C. Brambilla. 1996. Mécanismes de la carcinogenèse bronchique. *Encycl Med Chir*.
- Brambilla, E., D. Moro, S. Gazzeri, and C. Brambilla. 1999. Alterations of expression of Rb, p16(INK4A) and cyclin D1 in non-small cell lung carcinoma and their clinical significance. *The Journal of pathology*. 188:351-360.
- Brar, A.K., C.A. Kessler, and S. Handwerger. 2002. An Ets motif in the proximal decidual prolactin promoter is essential for basal gene expression. *Journal of molecular endocrinology*. 29:99-112.
- Brehm, A., G. Langst, J. Kehle, C.R. Clapier, A. Imhof, A. Eberharter, J. Muller, and P.B. Becker. 2000. dMi-2 and ISWI chromatin remodelling factors have distinct nucleosome binding and mobilization properties. *The EMBO journal*. 19:4332-4341.
- Brehm, A., E.A. Miska, D.J. McCance, J.L. Reid, A.J. Bannister, and T. Kouzarides. 1998. Retinoblastoma protein recruits histone deacetylase to repress transcription. *Nature*. 391:597-601.
- Bremnes, R.M., K. Al-Shibli, T. Donnem, R. Sirera, S. Al-Saad, S. Andersen, H. Stenvold, C. Camps, and L.T. Busund. 2011. The role of tumor-infiltrating immune cells and chronic inflammation at the tumor site on cancer development, progression, and prognosis: emphasis on non-small cell lung cancer. *Journal of thoracic oncology : official publication of the International Association for the Study of Lung Cancer*. 6:824-833.
- Brichard, V.G., C. ; Lehmann, F. ; Louahed, J. 2008. Use of MAGE A3-Protein D Fusion Antigen in Immunotherapy Combined with Surgery, Chemotherapy or Radiotherapy for the Treatment of Cancer Vol. n°US 20100008980 A1. B.-R.B. GLAXOSMITHKLINE BIOLOGICALS SA ; rue de l'Institut 89, editor.
- Briskin, C., S. Kaur, T.E. Chavarria, N. Binart, R.L. Sutherland, R.A. Weinberg, P.A. Kelly, and C.J. Ormandy. 1999. Prolactin controls mammary gland development via direct and indirect mechanisms. *Developmental biology*. 210:96-106.
- Brock, M.V., C.M. Hooker, E. Ota-Machida, Y. Han, M. Guo, S. Ames, S. Glockner, S. Piantadosi, E. Gabrielson, G. Pridham, K. Pelosky, S.A. Belinsky, S.C. Yang, S.B. Baylin, and J.G. Herman. 2008. DNA methylation markers and early recurrence in stage I lung cancer. *The New England journal of medicine*. 358:1118-1128.
- Brockman, J.L., M.D. Schroeder, and L.A. Schuler. 2002. PRL activates the cyclin D1 promoter via the Jak2/Stat pathway. *Molecular endocrinology*. 16:774-784.
- Brose, M.S., P. Volpe, M. Feldman, M. Kumar, I. Rishi, R. Guerrero, E. Einhorn, M. Herlyn, J. Minna, A. Nicholson, J.A. Roth, S.M. Albelda, H. Davies, C. Cox, G. Brignell, P. Stephens, P.A. Futreal, R. Wooster, M.R. Stratton, and B.L. Weber. 2002. BRAF and RAS mutations in human lung cancer and melanoma. *Cancer research*. 62:6997-7000.
- Brosh, R., and V. Rotter. 2009. When mutants gain new powers: news from the mutant p53 field. *Nature reviews. Cancer*. 9:701-713.
- Broutin, I., J.B. Jomain, E. Tallet, J. van Agthoven, B. Raynal, S. Hoos, B.B. Kragelund, P.A. Kelly, A. Ducruix, P. England, and V. Goffin. 2010. Crystal structure of an affinity-matured prolactin complexed to its dimerized receptor reveals the topology of hormone binding site 2. *The Journal of biological chemistry*. 285:8422-8433.
- Burbee, D.G., E. Forgacs, S. Zochbauer-Muller, L. Shivakumar, K. Fong, B. Gao, D. Randle, M. Kondo, A. Virmani, S. Bader, Y. Sekido, F. Latif, S. Milchgrub, S. Toyooka, A.F. Gazdar, M.I. Lerman, E. Zbarovsky, M. White, and J.D. Minna. 2001. Epigenetic inactivation of RASSF1A in lung and breast cancers and malignant phenotype suppression. *Journal of the National Cancer Institute*. 93:691-699.
- Bustin, S.A., V. Benes, J.A. Garson, J. Hellemans, J. Huggett, M. Kubista, R. Mueller, T. Nolan, M.W. Pfaffl, G.L. Shipley, J. Vandesompele, and C.T. Wittwer. 2009. The MIQE guidelines: minimum information for publication of quantitative real-time PCR experiments. *Clinical chemistry*. 55:611-622.
- Cambois, E., F. Meslé, and G. Pison. 2009 L'allongement de la vie et ses conséquences en France, Institut National d'Etudes Démographiques.
- Campling, B.G., I.R. Sarda, K.A. Baer, S.C. Pang, H.M. Baker, W.S. Lofters, and T.G. Flynn. 1995. Secretion of atrial natriuretic peptide and vasopressin by small cell lung cancer. *Cancer*. 75:2442-2451.
- Cao, R., L. Wang, H. Wang, L. Xia, H. Erdjument-Bromage, P. Tempst, R.S. Jones, and Y. Zhang. 2002. Role of histone H3 lysine 27 methylation in Polycomb-group silencing. *Science*. 298:1039-1043.

- Carew, J.S., F.J. Giles, and S.T. Nawrocki. 2008. Histone deacetylase inhibitors: mechanisms of cell death and promise in combination cancer therapy. *Cancer letters*. 269:7-17.
- Carey, M.S., S.T. 2000. Transcriptional Regulation in Eukaryotes: Concepts, Strategies and Techniques.
- Caron, C., C. Lestrat, S. Marsal, E. Escoffier, S. Curtet, V. Virolle, P. Barbry, A. Debernardi, C. Brambilla, E. Brambilla, S. Rousseaux, and S. Khochbin. 2010. Functional characterization of ATAD2 as a new cancer/testis factor and a predictor of poor prognosis in breast and lung cancers. *Oncogene*. 29:5171-5181.
- Castellano, S., D. Kuck, M. Viviano, J. Yoo, F. Lopez-Vallejo, P. Conti, L. Tamborini, A. Pinto, J.L. Medina-Franco, and G. Sbardella. 2011. Synthesis and biochemical evaluation of delta(2)-isoxazoline derivatives as DNA methyltransferase 1 inhibitors. *Journal of medicinal chemistry*. 54:7663-7677.
- Catelinois, O., A. Rogel, D. Laurier, S. Billon, D. Hemon, P. Verger, and M. Tirmarche. 2006. Lung cancer attributable to indoor radon exposure in france: impact of the risk models and uncertainty analysis. *Environmental health perspectives*. 114:1361-1366.
- Cesare, A.J., and R.R. Reddel. 2010. Alternative lengthening of telomeres: models, mechanisms and implications. *Nature reviews. Genetics*. 11:319-330.
- Chalmel, F., and M. Primig. 2008. The Annotation, Mapping, Expression and Network (AMEN) suite of tools for molecular systems biology. *BMC bioinformatics*. 9:86.
- Chang, Y.F., J.S. Imam, and M.F. Wilkinson. 2007. The nonsense-mediated decay RNA surveillance pathway. *Annual review of biochemistry*. 76:51-74.
- Chatterjee, N., D. Sinha, M. Lemma-Dechassa, S. Tan, M.A. Shogren-Knaak, and B. Bartholomew. 2011. Histone H3 tail acetylation modulates ATP-dependent remodeling through multiple mechanisms. *Nucleic acids research*. 39:8378-8391.
- Chavez-Rueda, K., J. Hernandez, E. Zenteno, A. Leanos-Miranda, M.V. Legorreta-Haquet, and F. Blanco-Favela. 2005. Identification of prolactin as a novel immunomodulator on the expression of co-stimulatory molecules and cytokine secretions on T and B human lymphocytes. *Clinical immunology*. 116:182-191.
- Chen, J.D., and R.M. Evans. 1995. A transcriptional co-repressor that interacts with nuclear hormone receptors. *Nature*. 377:454-457.
- Chen, W.Y., P. Ramamoorthy, N. Chen, R. Sticca, and T.E. Wagner. 1999. A human prolactin antagonist, hPRL-G129R, inhibits breast cancer cell proliferation through induction of apoptosis. *Clinical cancer research : an official journal of the American Association for Cancer Research*. 5:3583-3593.
- Chen, Y.T., M.J. Scanlan, U. Sahin, O. Tureci, A.O. Gure, S. Tsang, B. Williamson, E. Stockert, M. Pfreundschuh, and L.J. Old. 1997. A testicular antigen aberrantly expressed in human cancers detected by autologous antibody screening. *Proceedings of the National Academy of Sciences of the United States of America*. 94:1914-1918.
- Chen, Y.T., E. Stockert, Y. Chen, P. Garin-Chesa, W.J. Rettig, P. van der Bruggen, T. Boon, and L.J. Old. 1994. Identification of the MAGE-1 gene product by monoclonal and polyclonal antibodies. *Proceedings of the National Academy of Sciences of the United States of America*. 91:1004-1008.
- Chi, P., C.D. Allis, and G.G. Wang. 2010. Covalent histone modifications--miswritten, misinterpreted and mis-erased in human cancers. *Nature reviews. Cancer*. 10:457-469.
- Chilukamari, L., A.L. Hancock, S. Malik, J. Zabkiewicz, J.A. Baker, A. Greenhough, A.R. Dallosso, T.H. Huang, B. Royer-Pokora, K.W. Brown, and K. Malik. 2007. Hypomethylation and aberrant expression of the glioma pathogenesis-related 1 gene in Wilms tumors. *Neoplasia*. 9:970-978.
- Chiriva-Internati, M., A. Pandey, R. Saba, M. Kim, C. Saadeh, T. Lukman, R. Chiaramonte, M. Jenkins, E. Cobos, C. Jumper, and R. Alalawi. 2012. Cancer testis antigens: a novel target in lung cancer. *International reviews of immunology*. 31:321-343.
- Chiu, H.F., M.H. Cheng, S.S. Tsai, T.N. Wu, H.W. Kuo, and C.Y. Yang. 2006. Outdoor air pollution and female lung cancer in Taiwan. *Inhalation toxicology*. 18:1025-1031.
- Chiurazzi, P., M.G. Pomponi, R. Willemsen, B.A. Oostra, and G. Neri. 1998. In vitro reactivation of the FMR1 gene involved in fragile X syndrome. *Human molecular genetics*. 7:109-113.
- Choi, Y.H. 2005. Induction of apoptosis by trichostatin A, a histone deacetylase inhibitor, is associated with inhibition of cyclooxygenase-2 activity in human non-small cell lung cancer cells. *International journal of oncology*. 27:473-479.
- Clapp, C., J. Aranda, C. Gonzalez, M.C. Jeziorski, and G. Martinez de la Escalera. 2006. Vasoinhibins: endogenous regulators of angiogenesis and vascular function. *Trends in endocrinology and metabolism: TEM*. 17:301-307.
- Clapp, C., J.A. Martial, R.C. Guzman, F. Rentier-Delure, and R.I. Weiner. 1993. The 16-kilodalton N-terminal fragment of human prolactin is a potent inhibitor of angiogenesis. *Endocrinology*. 133:1292-1299.
- Clegg, A., D.A. Scott, P. Hewitson, M. Sidhu, and N. Waugh. 2002. Clinical and cost effectiveness of paclitaxel, docetaxel, gemcitabine, and vinorelbine in non-small cell lung cancer: a systematic review. *Thorax*. 57:20-28.
- Clemens, J.A., and C.J. Shaar. 1980. Control of prolactin secretion in mammals. *Federation proceedings*. 39:2588-2592.
- Clendenen, T.V., A.A. Arslan, A.E. Lokshin, M. Liu, E. Lundin, K.L. Koenig, F. Berrino, G. Hallmans, A. Idahl, V. Krogh, A. Lukanova, A. Marrangoni, P. Muti, B.M. Nolen, N. Ohlson, R.E. Shore, S. Sieri, and A. Zeleniuch-Jacquotte. 2013. Circulating prolactin levels and risk of epithelial ovarian cancer. *Cancer causes & control : CCC*. 24:741-748.
- Clevenger, C.V., W.P. Chang, W. Ngo, T.L. Pasha, K.T. Montone, and J.E. Tomaszewski. 1995. Expression of prolactin and prolactin receptor in human breast carcinoma. Evidence for an autocrine/paracrine loop. *The American journal of pathology*. 146:695-705.

- Clevenger, C.V., P.A. Furth, S.E. Hankinson, and L.A. Schuler. 2003. The role of prolactin in mammary carcinoma. *Endocrine reviews*. 24:1-27.
- Clevenger, C.V., and J.B. Kline. 2001. Prolactin receptor signal transduction. *Lupus*. 10:706-718.
- Cohen, L.E., F.E. Wondisford, and S. Radovick. 1996. Role of Pit-1 in the gene expression of growth hormone, prolactin, and thyrotropin. *Endocrinology and metabolism clinics of North America*. 25:523-540.
- Cohen, M.H., P.J. Creaven, B.E. Fossieck, Jr., L.E. Broder, O.S. Selawry, A.V. Johnston, C.L. Williams, and J.D. Minna. 1977. Intensive chemotherapy of small cell bronchogenic carcinoma. *Cancer treatment reports*. 61:349-354.
- Collins, K., and J.R. Mitchell. 2002. Telomerase in the human organism. *Oncogene*. 21:564-579.
- Collins, L.G., C. Haines, R. Perkel, and R.E. Enck. 2007. Lung cancer: diagnosis and management. *American family physician*. 75:56-63.
- Cooke, N.E., D. Coit, J. Shine, J.D. Baxter, and J.A. Martial. 1981. Human prolactin. cDNA structural analysis and evolutionary comparisons. *The Journal of biological chemistry*. 256:4007-4016.
- Coqueret, O. 2002. Linking cyclins to transcriptional control. *Gene*. 299:35-55.
- Corpet, F. 1988. Multiple sequence alignment with hierarchical clustering. *Nucleic acids research*. 16:10881-10890.
- Costello, J.F., M.C. Fruhwald, D.J. Smiraglia, L.J. Rush, G.P. Robertson, X. Gao, F.A. Wright, J.D. Feramisco, P. Peltomaki, J.C. Lang, D.E. Schuller, L. Yu, C.D. Bloomfield, M.A. Caligiuri, A. Yates, R. Nishikawa, H. Su Huang, N.J. Petrelli, X. Zhang, M.S. O'Dorisio, W.A. Held, W.K. Cavenee, and C. Plass. 2000. Aberrant CpG-island methylation has non-random and tumour-type-specific patterns. *Nature genetics*. 24:132-138.
- Counter, C.M., A.A. Avilion, C.E. LeFeuvre, N.G. Stewart, C.W. Greider, C.B. Harley, and S. Bacchetti. 1992. Telomere shortening associated with chromosome instability is arrested in immortal cells which express telomerase activity. *The EMBO journal*. 11:1921-1929.
- Cremer, T., and C. Cremer. 2001. Chromosome territories, nuclear architecture and gene regulation in mammalian cells. *Nature reviews. Genetics*. 2:292-301.
- D'Alessio, A.C., I.C. Weaver, and M. Szyf. 2007. Acetylation-induced transcription is required for active DNA demethylation in methylation-silenced genes. *Molecular and cellular biology*. 27:7462-7474.
- D'Angelo, G., J.F. Martini, T. Iiri, W.J. Fantl, J. Martial, and R.I. Weiner. 1999. 16K human prolactin inhibits vascular endothelial growth factor-induced activation of Ras in capillary endothelial cells. *Molecular endocrinology*. 13:692-704.
- D'Arcy, V., Z.K. Abdullaev, N. Pore, F. Docquier, V. Torrano, I. Chernukhin, M. Smart, D. Farrar, M. Metodiev, N. Fernandez, C. Richard, M.D. Delgado, V. Lobanekov, and E. Klenova. 2006. The potential of BORIS detected in the leukocytes of breast cancer patients as an early marker of tumorigenesis. *Clinical cancer research : an official journal of the American Association for Cancer Research*. 12:5978-5986.
- D'Arcy, V., N. Pore, F. Docquier, Z.K. Abdullaev, I. Chernukhin, G.X. Kita, S. Rai, M. Smart, D. Farrar, S. Pack, V. Lobanekov, and E. Klenova. 2008. BORIS, a paralogue of the transcription factor, CTCF, is aberrantly expressed in breast tumours. *British journal of cancer*. 98:571-579.
- D'Urso, V., V. Doneddu, I. Marchesi, A. Collodoro, P. Pirina, A. Giordano, and L. Bagella. 2013. Sputum analysis: non-invasive early lung cancer detection. *Journal of cellular physiology*. 228:945-951.
- Dacic, S. 2011. Molecular diagnostics of lung carcinomas. *Archives of pathology & laboratory medicine*. 135:622-629.
- Dagvadorj, A., S. Collins, J.B. Jomain, J. Abdulghani, J. Karras, T. Zellweger, H. Li, M. Nurmi, K. Alanen, T. Mirtti, T. Visakorpi, L. Bubendorf, V. Goffin, and M.T. Nevalainen. 2007. Autocrine prolactin promotes prostate cancer cell growth via Janus kinase-2-signal transducer and activator of transcription-5a/b signaling pathway. *Endocrinology*. 148:3089-3101.
- Damiano, J.S., and E. Wasserman. 2013. Molecular pathways: blockade of the PRLR signaling pathway as a novel antihormonal approach for the treatment of breast and prostate cancer. *Clinical cancer research : an official journal of the American Association for Cancer Research*. 19:1644-1650.
- Dannies, P.S. 2002. Mechanisms for storage of prolactin and growth hormone in secretory granules. *Molecular genetics and metabolism*. 76:6-13.
- Dannies, P.S. 2012. Prolactin and growth hormone aggregates in secretory granules: the need to understand the structure of the aggregate. *Endocrine reviews*. 33:254-270.
- Darst, R.P., D. Wang, and D.T. Auble. 2001. MOT1-catalyzed TBP-DNA disruption: uncoupling DNA conformational change and role of upstream DNA. *The EMBO journal*. 20:2028-2040.
- Dasanu, C.A., B.A. Clark, 3rd, B. Lahiri, T.E. Ichim, and D.T. Alexandrescu. 2010. Small cell lung cancer with paraneoplastic lipase production. *Southern medical journal*. 103:819-822.
- Davuluri, R.V., I. Grosse, and M.Q. Zhang. 2001. Computational identification of promoters and first exons in the human genome. *Nature genetics*. 29:412-417.
- De Backer, O., K.C. Arden, M. Boretti, V. Vantomme, C. De Smet, S. Czekay, C.S. Viars, E. De Plaen, F. Brasseur, P. Chomez, B. Van den Eynde, T. Boon, and P. van der Bruggen. 1999. Characterization of the GAGE genes that are expressed in various human cancers and in normal testis. *Cancer research*. 59:3157-3165.
- de Greef, W.J., J.L. Voogt, T.J. Visser, S.W. Lamberts, and P. van der Schoot. 1987. Control of prolactin release induced by suckling. *Endocrinology*. 121:316-322.
- de la Serna, I.L., Y. Ohkawa, and A.N. Imbalzano. 2006. Chromatin remodelling in mammalian differentiation: lessons from ATP-dependent remodellers. *Nature reviews. Genetics*. 7:461-473.

- de Mello, R.A., B.M. Costa, R.M. Reis, and V. Hespanhol. 2012. Insights into angiogenesis in non-small cell lung cancer: molecular mechanisms, polymorphic genes, and targeted therapies. *Recent patents on anti-cancer drug discovery*. 7:118-131.
- De Smet, C., A. Lorient, and T. Boon. 2004. Promoter-dependent mechanism leading to selective hypomethylation within the 5' region of gene MAGE-A1 in tumor cells. *Molecular and cellular biology*. 24:4781-4790.
- De Smet, C., C. Lurquin, B. Lethe, V. Martelange, and T. Boon. 1999. DNA methylation is the primary silencing mechanism for a set of germ line- and tumor-specific genes with a CpG-rich promoter. *Molecular and cellular biology*. 19:7327-7335.
- Defossez, P.A., and I. Stancheva. 2011. Biological functions of methyl-CpG-binding proteins. *Progress in molecular biology and translational science*. 101:377-398.
- Delaval, K., and R. Feil. 2004. Epigenetic regulation of mammalian genomic imprinting. *Current opinion in genetics & development*. 14:188-195.
- Denissenko, M.F., A. Pao, M. Tang, and G.P. Pfeifer. 1996. Preferential formation of benzo[a]pyrene adducts at lung cancer mutational hotspots in P53. *Science*. 274:430-432.
- Denslow, S.A., and P.A. Wade. 2007. The human Mi-2/NuRD complex and gene regulation. *Oncogene*. 26:5433-5438.
- Detich, N., J. Theberge, and M. Szyf. 2002. Promoter-specific activation and demethylation by MBD2/demethylase. *The Journal of biological chemistry*. 277:35791-35794.
- Devesa, S.S., F. Bray, A.P. Vizcaino, and D.M. Parkin. 2005. International lung cancer trends by histologic type: male:female differences diminishing and adenocarcinoma rates rising. *International journal of cancer. Journal international du cancer*. 117:294-299.
- Dey, A., F. Chitsaz, A. Abbasi, T. Misteli, and K. Ozato. 2003. The double bromodomain protein Brd4 binds to acetylated chromatin during interphase and mitosis. *Proceedings of the National Academy of Sciences of the United States of America*. 100:8758-8763.
- Dhalluin, C., J.E. Carlson, L. Zeng, C. He, A.K. Aggarwal, and M.M. Zhou. 1999. Structure and ligand of a histone acetyltransferase bromodomain. *Nature*. 399:491-496.
- Dillon, N. 2012. Factor mediated gene priming in pluripotent stem cells sets the stage for lineage specification. *BioEssays : news and reviews in molecular, cellular and developmental biology*. 34:194-204.
- DiMattia, G.E., B. Gellersen, H.G. Bohnet, and H.G. Friesen. 1988. A human B-lymphoblastoid cell line produces prolactin. *Endocrinology*. 122:2508-2517.
- DiMattia, G.E., B. Gellersen, M.L. Duckworth, and H.G. Friesen. 1990. Human prolactin gene expression. The use of an alternative noncoding exon in decidua and the IM-9-P3 lymphoblast cell line. *The Journal of biological chemistry*. 265:16412-16421.
- DiMattia, G.E., R. Varghese, and G.F. Wagner. 1998. Molecular cloning and characterization of stanniocalcin-related protein. *Molecular and cellular endocrinology*. 146:137-140.
- Ding, L., G. Getz, D.A. Wheeler, E.R. Mardis, M.D. McLellan, K. Cibulskis, C. Sougnez, H. Greulich, D.M. Muzny, M.B. Morgan, L. Fulton, R.S. Fulton, Q. Zhang, M.C. Wendl, M.S. Lawrence, D.E. Larson, K. Chen, D.J. Dooling, A. Sabo, A.C. Hawes, H. Shen, S.N. Jhangiani, L.R. Lewis, O. Hall, Y. Zhu, T. Mathew, Y. Ren, J. Yao, S.E. Scherer, K. Clerc, G.A. Metcalf, B. Ng, A. Milosavljevic, M.L. Gonzalez-Garay, J.R. Osborne, R. Meyer, X. Shi, Y. Tang, D.C. Koboldt, L. Lin, R. Abbott, T.L. Miner, C. Pohl, G. Fewell, C. Haipek, H. Schmidt, B.H. Dunford-Shore, A. Kraja, S.D. Crosby, C.S. Sawyer, T. Vickery, S. Sander, J. Robinson, W. Winckler, J. Baldwin, L.R. Chirieac, A. Dutt, T. Fennell, M. Hanna, B.E. Johnson, R.C. Onofrio, R.K. Thomas, G. Tonon, B.A. Weir, X. Zhao, L. Ziaugra, M.C. Zody, T. Giordano, M.B. Orringer, J.A. Roth, M.R. Spitz, Wistuba, II, B. Ozenberger, P.J. Good, A.C. Chang, D.G. Beer, M.A. Watson, M. Ladanyi, S. Broderick, A. Yoshizawa, W.D. Travis, W. Pao, M.A. Province, G.M. Weinstock, H.E. Varmus, S.B. Gabriel, E.S. Lander, R.A. Gibbs, M. Meyerson, and R.K. Wilson. 2008. Somatic mutations affect key pathways in lung adenocarcinoma. *Nature*. 455:1069-1075.
- Divine, K.K., L.C. Pulling, P.G. Marron-Terada, K.C. Liechty, T. Kang, A.G. Schwartz, T.J. Bocklage, T.A. Coons, F.D. Gilliland, and S.A. Belinsky. 2005. Multiplicity of abnormal promoter methylation in lung adenocarcinomas from smokers and never smokers. *International journal of cancer. Journal international du cancer*. 114:400-405.
- Doll, R., and A.B. Hill. 1956. Lung cancer and other causes of death in relation to smoking; a second report on the mortality of British doctors. *British medical journal*. 2:1071-1081.
- Dong, J., S. Sutor, G. Jiang, Y. Cao, Y.W. Asmann, and D.A. Wigle. 2011. c-Myc regulates self-renewal in bronchoalveolar stem cells. *PloS one*. 6:e23707.
- Duan, Z., Y. Duan, D.E. Lamendola, R.Z. Yusuf, R. Naeem, R.T. Penson, and M.V. Seiden. 2003. Overexpression of MAGE/GAGE genes in paclitaxel/doxorubicin-resistant human cancer cell lines. *Clinical cancer research : an official journal of the American Association for Cancer Research*. 9:2778-2785.
- Dunn, B.K. 2003. Hypomethylation: one side of a larger picture. *Annals of the New York Academy of Sciences*. 983:28-42.
- Duvic, M., and J. Vu. 2007. Vorinostat: a new oral histone deacetylase inhibitor approved for cutaneous T-cell lymphoma. *Expert opinion on investigational drugs*. 16:1111-1120.
- Eberharter, A., and P.B. Becker. 2002. Histone acetylation: a switch between repressive and permissive chromatin. Second in review series on chromatin dynamics. *EMBO reports*. 3:224-229.
- Eden, A., F. Gaudet, A. Waghmare, and R. Jaenisch. 2003. Chromosomal instability and tumors promoted by DNA hypomethylation. *Science*. 300:455.

- El-Nikhely, N., L. Larzabal, W. Seeger, A. Calvo, and R. Savai. 2012. Tumor-stromal interactions in lung cancer: novel candidate targets for therapeutic intervention. *Expert opinion on investigational drugs*. 21:1107-1122.
- Emadali, A.R., S. ; Bruder-Costa, J. ; Rome, C. ; Duley, S. ; Hamaidia, S. ; Betton, P. ; Debernardi, A. ; Leroux, D. ; Bernay, B. ; Kieffer-Jaquinod, S. ; Combes, F. ; Ferri, E. ; McKenna, C.E. ; Petosa, C. ; Bruley, C. ; Garin, J. ; Ferro, M. ; Gressin, R. ; Callanan, M. B ; Khochbin, S. . 2013. Identification of a novel BET bromodomain inhibitor-sensitive, gene regulatory circuit that controls Rituximab response and tumor growth in aggressive lymphoid cancers. *EMBO Mol Med*.
- Engelman, J.A., K. Zejnullahu, T. Mitsudomi, Y. Song, C. Hyland, J.O. Park, N. Lindeman, C.M. Gale, X. Zhao, J. Christensen, T. Kosaka, A.J. Holmes, A.M. Rogers, F. Cappuzzo, T. Mok, C. Lee, B.E. Johnson, L.C. Cantley, and P.A. Janne. 2007. MET amplification leads to gefitinib resistance in lung cancer by activating ERBB3 signaling. *Science*. 316:1039-1043.
- Esteller, M. 2007. Epigenetic gene silencing in cancer: the DNA hypermethylome. *Human molecular genetics*. 16 Spec No 1:R50-59.
- Esteller, M. 2008. Epigenetics in cancer. *The New England journal of medicine*. 358:1148-1159.
- Esteller, M., and J.G. Herman. 2002. Cancer as an epigenetic disease: DNA methylation and chromatin alterations in human tumours. *The Journal of pathology*. 196:1-7.
- Esteller, M., M. Sanchez-Cespedes, R. Rosell, D. Sidransky, S.B. Baylin, and J.G. Herman. 1999. Detection of aberrant promoter hypermethylation of tumor suppressor genes in serum DNA from non-small cell lung cancer patients. *Cancer research*. 59:67-70.
- Evans, A.M., J.W. Petersen, G.S. Sekhon, and R. DeMars. 1989. Mapping of prolactin and tumor necrosis factor-beta genes on human chromosome 6p using lymphoblastoid cell deletion mutants. *Somatic cell and molecular genetics*. 15:203-213.
- Fanti, L., M. Berloco, L. Piacentini, and S. Pimpinelli. 2003. Chromosomal distribution of heterochromatin protein 1 (HP1) in Drosophila: a cytological map of euchromatic HP1 binding sites. *Genetica*. 117:135-147.
- Featherstone, K., M.R. White, and J.R. Davis. 2012. The prolactin gene: a paradigm of tissue-specific gene regulation with complex temporal transcription dynamics. *Journal of neuroendocrinology*. 24:977-990.
- Fedi, P., S.R. Tronick, S.A. Aaronson, J.F. Holland, R.C. Bast, D.L. Morton, E. Frei, D.W. Kufe, and W. R.R. 1997. Cancer Medicine. 41-64 pp.
- Feinberg, A.P., C.W. Gehrke, K.C. Kuo, and M. Ehrlich. 1988. Reduced genomic 5-methylcytosine content in human colonic neoplasia. *Cancer research*. 48:1159-1161.
- Feinberg, A.P., and B. Tycko. 2004. The history of cancer epigenetics. *Nature reviews. Cancer*. 4:143-153.
- Feld, R., S.S. Sridhar, F.A. Shepherd, J.A. Mackay, W.K. Evans, and C. Lung Cancer Disease Site Group of Cancer Care Ontario's Program in Evidence-based. 2006. Use of the epidermal growth factor receptor inhibitors gefitinib and erlotinib in the treatment of non-small cell lung cancer: a systematic review. *Journal of thoracic oncology : official publication of the International Association for the Study of Lung Cancer*. 1:367-376.
- Felsenfeld, G., J. Boyes, J. Chung, D. Clark, and V. Studitsky. 1996. Chromatin structure and gene expression. *Proceedings of the National Academy of Sciences of the United States of America*. 93:9384-9388.
- Fennell, D.A. 2003. Bcl-2 as a target for overcoming chemoresistance in small-cell lung cancer. *Clinical lung cancer*. 4:307-313.
- Ferlay, J., H.R. Shin, F. Bray, D. Forman, C. Mathers, and D.M. Parkin. 2010. Estimates of worldwide burden of cancer in 2008: GLOBOCAN 2008. *International journal of cancer. Journal international du cancer*. 127:2893-2917.
- Fernandez, I., P. Touraine, and V. Goffin. 2010. Prolactin and human tumorigenesis. *Journal of neuroendocrinology*. 22:771-777.
- Fields, K., E. Kulig, and R.V. Lloyd. 1993. Detection of prolactin messenger RNA in mammary and other normal and neoplastic tissues by polymerase chain reaction. *Laboratory investigation; a journal of technical methods and pathology*. 68:354-360.
- Fiering, S., E. Whitelaw, and D.I. Martin. 2000. To be or not to be active: the stochastic nature of enhancer action. *BioEssays : news and reviews in molecular, cellular and developmental biology*. 22:381-387.
- Fijak, M., S. Bhushan, and A. Meinhardt. 2011. Immunoprivileged sites: the testis. *Methods in molecular biology*. 677:459-470.
- Filippakopoulos, P., J. Qi, S. Picaud, Y. Shen, W.B. Smith, O. Fedorov, E.M. Morse, T. Keates, T.T. Hickman, I. Felletar, M. Philpott, S. Munro, M.R. McKeown, Y. Wang, A.L. Christie, N. West, M.J. Cameron, B. Schwartz, T.D. Heightman, N. La Thangue, C.A. French, O. Wiest, A.L. Kung, S. Knapp, and J.E. Bradner. 2010. Selective inhibition of BET bromodomains. *Nature*. 468:1067-1073.
- Finch, J.T., and A. Klug. 1976. Solenoidal model for superstructure in chromatin. *Proceedings of the National Academy of Sciences of the United States of America*. 73:1897-1901.
- Finch, J.T., L.C. Lutter, D. Rhodes, R.S. Brown, B. Rushton, M. Levitt, and A. Klug. 1977. Structure of nucleosome core particles of chromatin. *Nature*. 269:29-36.
- Fishman, J., D. Fukushima, J. O'Connor, R.S. Rosenfeld, H.T. Lynch, J.F. Lynch, H. Guirgis, and K. Maloney. 1978. Plasma hormone profiles of young women at risk for familial breast cancer. *Cancer research*. 38:4006-4011.
- Foulds, L. 1954. The Experimental Study of Tumor Progression. London: Academic Press.
- Freeman, M.E., B. Kanyicska, A. Lerant, and G. Nagy. 2000. Prolactin: structure, function, and regulation of secretion. *Physiological reviews*. 80:1523-1631.

- Fresno Vara, J.A., M.A. Caceres, A. Silva, and J. Martin-Perez. 2001. Src family kinases are required for prolactin induction of cell proliferation. *Molecular biology of the cell*. 12:2171-2183.
- Fritz, A., C. Percy, A. Jack, K. Shanmugaratnam, L. Sobin, M. Parkin, and S. Whelan. 2008. Classification internationale des maladies pour l'oncologie.
- Fukazawa, H., and A. Masumi. 2012. The conserved 12-amino acid stretch in the inter-bromodomain region of BET family proteins functions as a nuclear localization signal. *Biological & pharmaceutical bulletin*. 35:2064-2068.
- Fuks, F. 2003. [DNA methyltransferases: from chromatin remodeling to cancer]. *Medecine sciences : M/S*. 19:477-480.
- Fyodorov, D.V., and J.T. Kadonaga. 2001. The many faces of chromatin remodeling: SWItching beyond transcription. *Cell*. 106:523-525.
- Gangaraju, V.K., and B. Bartholomew. 2007. Mechanisms of ATP dependent chromatin remodeling. *Mutation research*. 618:3-17.
- Gansner, J.M., and I.O. Rosas. 2013. Telomeres in lung disease. *Translational research : the journal of laboratory and clinical medicine*.
- Garcia-Gutierrez, P., M. Mundi, and M. Garcia-Dominguez. 2012. Association of bromodomain BET proteins with chromatin requires dimerization through the conserved motif B. *Journal of cell science*. 125:3671-3680.
- Gardiner-Garden, M., and M. Frommer. 1987. CpG islands in vertebrate genomes. *Journal of molecular biology*. 196:261-282.
- Gaucher, J., F. Boussouar, E. Montellier, S. Curtet, T. Buchou, S. Bertrand, P. Hery, S. Jounier, A. Depaux, A.L. Vitte, P. Guardiola, K. Pernet, A. Debernardi, F. Lopez, H. Holota, J. Imbert, D.J. Wolgemuth, M. Gerard, S. Rousseaux, and S. Khochbin. 2012. Bromodomain-dependent stage-specific male genome programming by Brdt. *The EMBO journal*. 31:3809-3820.
- Gaudet, F., J.G. Hodgson, A. Eden, L. Jackson-Grusby, J. Dausman, J.W. Gray, H. Leonhardt, and R. Jaenisch. 2003. Induction of tumors in mice by genomic hypomethylation. *Science*. 300:489-492.
- Gautschi, O., D. Ratschiller, M. Gugger, D.C. Betticher, and J. Heighway. 2007. Cyclin D1 in non-small cell lung cancer: a key driver of malignant transformation. *Lung cancer*. 55:1-14.
- Gazzeri, S., V. Gouyer, C. Vour'ch, C. Brambilla, and E. Brambilla. 1998. Mechanisms of p16INK4A inactivation in non small-cell lung cancers. *Oncogene*. 16:497-504.
- Gellersen, B., G.E. DiMattia, H.G. Friesen, and H.G. Bohnet. 1989. Prolactin (PRL) mRNA from human decidua differs from pituitary PRL mRNA but resembles the IM-9-P3 lymphoblast PRL transcript. *Molecular and cellular endocrinology*. 64:127-130.
- Gellersen, B., R. Kempf, R. Telgmann, and G.E. DiMattia. 1994. Nonpituitary human prolactin gene transcription is independent of Pit-1 and differentially controlled in lymphocytes and in endometrial stroma. *Molecular endocrinology*. 8:356-373.
- Gerlo, S., J.R. Davis, D.L. Mager, and R. Kooijman. 2006. Prolactin in man: a tale of two promoters. *BioEssays : news and reviews in molecular, cellular and developmental biology*. 28:1051-1055.
- Ghafouri-Fard, S., Z. Ousati Ashtiani, B. Sabah Golian, S.M. Hasheminasab, and M.H. Modarressi. 2010. Expression of two testis-specific genes, SPATA19 and LEMD1, in prostate cancer. *Archives of medical research*. 41:195-200.
- Goffin, J., and E. Eisenhauer. 2002. DNA methyltransferase inhibitors-state of the art. *Annals of oncology : official journal of the European Society for Medical Oncology / ESMO*. 13:1699-1716.
- Goffin, V., N. Binart, P. Touraine, and P.A. Kelly. 2002. Prolactin: the new biology of an old hormone. *Annual review of physiology*. 64:47-67.
- Goffin, V., D.T. Hoang, R.L. Bogorad, and M.T. Nevalainen. 2011. Prolactin regulation of the prostate gland: a female player in a male game. *Nature reviews. Urology*. 8:597-607.
- Goffin, V., P. Touraine, M.D. Culler, and P.A. Kelly. 2006. Drug Insight: prolactin-receptor antagonists, a novel approach to treatment of unresolved systemic and local hyperprolactinemia? *Nature clinical practice. Endocrinology & metabolism*. 2:571-581.
- Gokul, G., B. Gautami, S. Malathi, A.P. Sowjanya, U.R. Poli, M. Jain, G. Ramakrishna, and S. Khosla. 2007. DNA methylation profile at the DNMT3L promoter: a potential biomarker for cervical cancer. *Epigenetics : official journal of the DNA Methylation Society*. 2:80-85.
- Goldstraw, P., J. Crowley, K. Chansky, D.J. Giroux, P.A. Groome, R. Rami-Porta, P.E. Postmus, V. Rusch, L. Sobin, C. International Association for the Study of Lung Cancer International Staging, and I. Participating. 2007. The IASLC Lung Cancer Staging Project: proposals for the revision of the TNM stage groupings in the forthcoming (seventh) edition of the TNM Classification of malignant tumours. *Journal of thoracic oncology : official publication of the International Association for the Study of Lung Cancer*. 2:706-714.
- Goll, M.G., and T.H. Bestor. 2005. Eukaryotic cytosine methyltransferases. *Annual review of biochemistry*. 74:481-514.
- Gomperts, B.N., A. Spira, P.P. Massion, T.C. Walser, Wistuba, II, J.D. Minna, and S.M. Dubinett. 2011. Evolving concepts in lung carcinogenesis. *Seminars in respiratory and critical care medicine*. 32:32-43.
- Goodman, R.H., and S. Smolik. 2000. CBP/p300 in cell growth, transformation, and development. *Genes & development*. 14:1553-1577.
- Goodsell, D. 2011. Dna Methyltransferases, Molecules of the Month. *PDB Protein Data Bank*. July 2011.
- Gorlova, O.Y., S.F. Weng, Y. Zhang, C.I. Amos, and M.R. Spitz. 2007. Aggregation of cancer among relatives of never-smoking lung cancer patients. *International journal of cancer. Journal international du cancer*. 121:111-118.

- Gouilleux, F., H. Wakao, M. Mundt, and B. Groner. 1994. Prolactin induces phosphorylation of Tyr694 of Stat5 (MGF), a prerequisite for DNA binding and induction of transcription. *The EMBO journal*. 13:4361-4369.
- Gouyer, V., S. Gazzeri, E. Brambilla, I. Bolon, D. Moro, P. Perron, A.L. Benabid, and C. Brambilla. 1994. Loss of heterozygosity at the RB locus correlates with loss of RB protein in primary malignant neuro-endocrine lung carcinomas. *International journal of cancer. Journal international du cancer*. 58:818-824.
- Govin, J., C. Caron, E. Escoffier, M. Ferro, L. Kuhn, S. Rousseaux, E.M. Eddy, J. Garin, and S. Khochbin. 2006. Post-meiotic shifts in HSPA2/HSP70.2 chaperone activity during mouse spermatogenesis. *The Journal of biological chemistry*. 281:37888-37892.
- Govin, J., C. Caron, C. Lestrat, S. Rousseaux, and S. Khochbin. 2004. The role of histones in chromatin remodelling during mammalian spermiogenesis. *European journal of biochemistry / FEBS*. 271:3459-3469.
- Govin, J., E. Escoffier, S. Rousseaux, L. Kuhn, M. Ferro, J. Thevenon, R. Catena, I. Davidson, J. Garin, S. Khochbin, and C. Caron. 2007. Pericentric heterochromatin reprogramming by new histone variants during mouse spermiogenesis. *The Journal of cell biology*. 176:283-294.
- Grah, J., M. Samija, A. Juretic, B. Sarcevic, and H. Sobat. 2008. Immunohistochemical expression of cancer/testis antigens (MAGE-A3/4, NY-ESO-1) in non-small cell lung cancer: the relationship with clinical-pathological features. *Collegium antropologicum*. 32:731-736.
- Grattan, D.R., and I.C. Kokay. 2008. Prolactin: a pleiotropic neuroendocrine hormone. *Journal of neuroendocrinology*. 20:752-763.
- Greenman, C., P. Stephens, R. Smith, G.L. Dalgliesh, C. Hunter, G. Bignell, H. Davies, J. Teague, A. Butler, C. Stevens, S. Edkins, S. O'Meara, I. Vastrik, E.E. Schmidt, T. Avis, S. Barthorpe, G. Bhamra, G. Buck, B. Choudhury, J. Clements, J. Cole, E. Dicks, S. Forbes, K. Gray, K. Halliday, R. Harrison, K. Hills, J. Hinton, A. Jenkinson, D. Jones, A. Menzies, T. Mironenko, J. Perry, K. Raine, D. Richardson, R. Shepherd, A. Small, C. Tofts, J. Varian, T. Webb, S. West, S. Widaa, A. Yates, D.P. Cahill, D.N. Louis, P. Goldstraw, A.G. Nicholson, F. Brasseur, L. Looijenga, B.L. Weber, Y.E. Chiew, A. DeFazio, M.F. Greaves, A.R. Green, P. Campbell, E. Birney, D.F. Easton, G. Chenevix-Trench, M.H. Tan, S.K. Khoo, B.T. Teh, S.T. Yuen, S.Y. Leung, R. Wooster, P.A. Futreal, and M.R. Stratton. 2007. Patterns of somatic mutation in human cancer genomes. *Nature*. 446:153-158.
- Greulich, H., T.H. Chen, W. Feng, P.A. Janne, J.V. Alvarez, M. Zappaterra, S.E. Bulmer, D.A. Frank, W.C. Hahn, W.R. Sellers, and M. Meyerson. 2005. Oncogenic transformation by inhibitor-sensitive and -resistant EGFR mutants. *PLoS medicine*. 2:e313.
- Grewal, S.I., and S. Jia. 2007. Heterochromatin revisited. *Nature reviews. Genetics*. 8:35-46.
- Grewal, S.I., and D. Moazed. 2003. Heterochromatin and epigenetic control of gene expression. *Science*. 301:798-802.
- Gridelli, C., A. Rossi, and P. Maione. 2008. The potential role of histone deacetylase inhibitors in the treatment of non-small-cell lung cancer. *Critical reviews in oncology/hematology*. 68:29-36.
- Grigorova, M., R.C. Lyman, C. Caldas, and P.A. Edwards. 2005. Chromosome abnormalities in 10 lung cancer cell lines of the NCI-H series analyzed with spectral karyotyping. *Cancer genetics and cytogenetics*. 162:1-9.
- Groner, A.C., S. Meylan, A. Ciuffi, N. Zangger, G. Ambrosini, N. Denervaud, P. Bucher, and D. Trono. 2010. KRAB-zinc finger proteins and KAP1 can mediate long-range transcriptional repression through heterochromatin spreading. *PLoS genetics*. 6:e1000869.
- Groner, B., and F. Gouilleux. 1995. Prolactin-mediated gene activation in mammary epithelial cells. *Current opinion in genetics & development*. 5:587-594.
- Gropp, C., K. Havemann, and A. Scheuer. 1980. Ectopic hormones in lung cancer patients at diagnosis and during therapy. *Cancer*. 46:347-354.
- Grosclaude, P., L. Remontet, A. Belot, A. Danzon, C. Rasamimanana, and N. Bossard. 2013. Survie des personnes atteintes de cancer en France, 1989-2007 - Etude à partir des registres des cancers du réseau Francim. Institut de veille sanitaire.
- Gross, A. 2001. BCL-2 proteins: regulators of the mitochondrial apoptotic program. *IUBMB life*. 52:231-236.
- Groudine, M., and H. Weintraub. 1981. Activation of globin genes during chicken development. *Cell*. 24:393-401.
- Grunstein, M. 1997. Histone acetylation in chromatin structure and transcription. *Nature*. 389:349-352.
- Gu, W., and R.G. Roeder. 1997. Activation of p53 sequence-specific DNA binding by acetylation of the p53 C-terminal domain. *Cell*. 90:595-606.
- Guerin, S., and C. Hill. 2010. [Cancer epidemiology in France in 2010, comparison with the USA]. *Bulletin du cancer*. 97:47-54.
- Gurdon, J.B., T.R. Eldsle, and M. Fischberg. 1958. Sexually mature individuals of *Xenopus laevis* from the transplantation of single somatic nuclei. *Nature*. 182:64-65.
- Gure, A.O., R. Chua, B. Williamson, M. Gonen, C.A. Ferrera, S. Gnjjatic, G. Ritter, A.J. Simpson, Y.T. Chen, L.J. Old, and N.K. Altorki. 2005. Cancer-testis genes are coordinately expressed and are markers of poor outcome in non-small cell lung cancer. *Clinical cancer research : an official journal of the American Association for Cancer Research*. 11:8055-8062.
- Gustafson, A.M., R. Soldi, C. Anderlind, M.B. Scholand, J. Qian, X. Zhang, K. Cooper, D. Walker, A. McWilliams, G. Liu, E. Szabo, J. Brody, P.P. Massion, M.E. Lenburg, S. Lam, A.H. Bild, and A. Spira. 2010. Airway PI3K pathway activation is an early and reversible event in lung cancer development. *Science translational medicine*. 2:26ra25.
- Guyette, W.A., R.J. Matusik, and J.M. Rosen. 1979. Prolactin-mediated transcriptional and post-transcriptional control of casein gene expression. *Cell*. 17:1013-1023.

- Hackshaw, A.K., M.R. Law, and N.J. Wald. 1997. The accumulated evidence on lung cancer and environmental tobacco smoke. *Bmj*. 315:980-988.
- Hainaut, P., and G.P. Pfeifer. 2001. Patterns of p53 G→T transversions in lung cancers reflect the primary mutagenic signature of DNA-damage by tobacco smoke. *Carcinogenesis*. 22:367-374.
- Hair, W.M., O. Gubbay, H.N. Jabbour, and G.A. Lincoln. 2002. Prolactin receptor expression in human testis and accessory tissues: localization and function. *Molecular human reproduction*. 8:606-611.
- Han, S.W., and J. Roman. 2010. Targeting apoptotic signaling pathways in human lung cancer. *Current cancer drug targets*. 10:566-574.
- Hanahan, D., and R.A. Weinberg. 2000. The hallmarks of cancer. *Cell*. 100:57-70.
- Hanahan, D., and R.A. Weinberg. 2011. Hallmarks of cancer: the next generation. *Cell*. 144:646-674.
- Hang, B. 2010. Formation and repair of tobacco carcinogen-derived bulky DNA adducts. *Journal of nucleic acids*. 2010:709521.
- Hankinson, S.E., W.C. Willett, D.S. Michaud, J.E. Manson, G.A. Colditz, C. Longcope, B. Rosner, and F.E. Speizer. 1999. Plasma prolactin levels and subsequent risk of breast cancer in postmenopausal women. *Journal of the National Cancer Institute*. 91:629-634.
- Harley, C.B., and S.W. Sherwood. 1997. Telomerase, checkpoints and cancer. *Cancer surveys*. 29:263-284.
- Harper, J.W., G.R. Adami, N. Wei, K. Keyomarsi, and S.J. Elledge. 1993. The p21 Cdk-interacting protein Cip1 is a potent inhibitor of G1 cyclin-dependent kinases. *Cell*. 75:805-816.
- Harris, R.A., T. Wang, C. Coarfa, R.P. Nagarajan, C. Hong, S.L. Downey, B.E. Johnson, S.D. Fouse, A. Delaney, Y. Zhao, A. Olshen, T. Ballinger, X. Zhou, K.J. Forsberg, J. Gu, L. Echipare, H. O'Geen, R. Lister, M. Pelizzola, Y. Xi, C.B. Epstein, B.E. Bernstein, R.D. Hawkins, B. Ren, W.Y. Chung, H. Gu, C. Bock, A. Gnirke, M.Q. Zhang, D. Haussler, J.R. Ecker, W. Li, P.J. Farnham, R.A. Waterland, A. Meissner, M.A. Marra, M. Hirst, A. Milosavljevic, and J.F. Costello. 2010. Comparison of sequencing-based methods to profile DNA methylation and identification of monoallelic epigenetic modifications. *Nature biotechnology*. 28:1097-1105.
- Hassan, A.H., K.E. Neely, and J.L. Workman. 2001. Histone acetyltransferase complexes stabilize swi/snf binding to promoter nucleosomes. *Cell*. 104:817-827.
- Hassan, A.H., P. Prochasson, K.E. Neely, S.C. Galasinski, M. Chandy, M.J. Carrozza, and J.L. Workman. 2002. Function and selectivity of bromodomains in anchoring chromatin-modifying complexes to promoter nucleosomes. *Cell*. 111:369-379.
- Hayflick, L. 1997. Mortality and immortality at the cellular level. A review. *Biochemistry. Biokhimiia*. 62:1180-1190.
- Hecht, S.S. 2003. Tobacco carcinogens, their biomarkers and tobacco-induced cancer. *Nature reviews. Cancer*. 3:733-744.
- Hecht, S.S. 2012. Lung carcinogenesis by tobacco smoke. *International journal of cancer. Journal international du cancer*. 131:2724-2732.
- Heintz, N.H., Y.M. Janssen-Heininger, and B.T. Mossman. 2010. Asbestos, lung cancers, and mesotheliomas: from molecular approaches to targeting tumor survival pathways. *American journal of respiratory cell and molecular biology*. 42:133-139.
- Heitz, E. 1928. Das Heterochromatin des Moose. *Jahrb Wiss Botanik* 69:762-818.
- Heller, G., C.C. Zielinski, and S. Zochbauer-Muller. 2010. Lung cancer: from single-gene methylation to methylome profiling. *Cancer metastasis reviews*. 29:95-107.
- Herman, J.G., and S.B. Baylin. 2003. Gene silencing in cancer in association with promoter hypermethylation. *The New England journal of medicine*. 349:2042-2054.
- Herman, J.G., F. Latif, Y. Weng, M.I. Lerman, B. Zbar, S. Liu, D. Samid, D.S. Duan, J.R. Gnarr, W.M. Linehan, and et al. 1994. Silencing of the VHL tumor-suppressor gene by DNA methylation in renal carcinoma. *Proceedings of the National Academy of Sciences of the United States of America*. 91:9700-9704.
- Hewish, D.R., and L.A. Burgoyne. 1973. Chromatin sub-structure. The digestion of chromatin DNA at regularly spaced sites by a nuclear deoxyribonuclease. *Biochemical and biophysical research communications*. 52:504-510.
- Hilfiker-Kleiner, D., K. Kaminski, E. Podewski, T. Bonda, A. Schaefer, K. Sliwa, O. Forster, A. Quint, U. Landmesser, C. Doerries, M. Luchtefeld, V. Poli, M.D. Schneider, J.L. Balligand, F. Desjardins, A. Ansari, I. Struman, N.Q. Nguyen, N.H. Zschemisch, G. Klein, G. Heusch, R. Schulz, A. Hilfiker, and H. Drexler. 2007. A cathepsin D-cleaved 16 kDa form of prolactin mediates postpartum cardiomyopathy. *Cell*. 128:589-600.
- Hiret, S., H. Senellart, and J. Bennouna. 2010. [Molecular biology of lung cancer series]. *Revue des maladies respiratoires*. 27:954-958.
- Hirose, Y., and Y. Ohkuma. 2007. Phosphorylation of the C-terminal domain of RNA polymerase II plays central roles in the integrated events of eucaryotic gene expression. *Journal of biochemistry*. 141:601-608.
- Hirsch, F.R., M. Varella-Garcia, P.A. Bunn, Jr., M.V. Di Maria, R. Veve, R.M. Bremmes, A.E. Baron, C. Zeng, and W.A. Franklin. 2003. Epidermal growth factor receptor in non-small-cell lung carcinomas: correlation between gene copy number and protein expression and impact on prognosis. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology*. 21:3798-3807.
- Hollstein, M., D. Sidransky, B. Vogelstein, and C.C. Harris. 1991. p53 mutations in human cancers. *Science*. 253:49-53.
- Holmquist, G.P., and S. Gao. 1997. Somatic mutation theory, DNA repair rates, and the molecular epidemiology of p53 mutations. *Mutation research*. 386:69-101.
- Hong, J.A., Y. Kang, Z. Abdullaev, P.T. Flanagan, S.D. Pack, M.R. Fischette, M.T. Adnani, D.I. Loukinov, S. Vatolin, J.I. Risinger, M. Custer, G.A. Chen, M. Zhao, D.M. Nguyen, J.C. Barrett, V.V. Lobanenko, and D.S. Schrupp. 2005.

- Reciprocal binding of CTCF and BORIS to the NY-ESO-1 promoter coincides with derepression of this cancer-testis gene in lung cancer cells. *Cancer research*. 65:7763-7774.
- Horlein, A.J., A.M. Naar, T. Heinzel, J. Torchia, B. Gloss, R. Kurokawa, A. Ryan, Y. Kamei, M. Soderstrom, C.K. Glass, and et al. 1995. Ligand-independent repression by the thyroid hormone receptor mediated by a nuclear receptor co-repressor. *Nature*. 377:397-404.
- Horn, L., and D.H. Johnson. 2008. Evarts A. Graham and the first pneumonectomy for lung cancer. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology*. 26:3268-3275.
- Horseman, N.D. 2002. Prolactin receptor diversity in humans: novel isoforms suggest general principles. *Trends in endocrinology and metabolism: TEM*. 13:47-48.
- Hsia, T.C., C.J. Liu, C.C. Chu, L.W. Hang, W.S. Chang, C.W. Tsai, C.I. Wu, C.S. Lien, W.L. Liao, C.Y. Ho, and D.T. Bau. 2012. Association of DNA double-strand break gene XRCC6 genotypes and lung cancer in Taiwan. *Anticancer research*. 32:1015-1020.
- Hu, W., Z. Feng, and M.S. Tang. 2003. Preferential carcinogen-DNA adduct formation at codons 12 and 14 in the human K-ras gene and their possible mechanisms. *Biochemistry*. 42:10012-10023.
- Huang, X., J. Li, L. Lu, M. Xu, J. Xiao, L. Yin, H. Zhu, Z. Zhou, and J. Sha. 2005. Novel development-related alternative splices in human testis identified by cDNA microarrays. *Journal of andrology*. 26:189-196.
- Huang, Y.T., X. Lin, Y. Liu, L.R. Chirieac, R. McGovern, J. Wain, R. Heist, V. Skaug, S. Zienolddiny, A. Haugen, L. Su, E.A. Fox, K.K. Wong, and D.C. Christiani. 2011. Cigarette smoking increases copy number alterations in nonsmall-cell lung cancer. *Proceedings of the National Academy of Sciences of the United States of America*. 108:16345-16350.
- Hubaux, R., D.D. Becker-Santos, K.S. Enfield, S. Lam, W.L. Lam, and V.D. Martinez. 2012. Arsenic, asbestos and radon: emerging players in lung tumorigenesis. *Environmental health : a global access science source*. 11:89.
- Huebner, R.J., and G.J. Todaro. 1969. Oncogenes of RNA tumor viruses as determinants of cancer. *Proceedings of the National Academy of Sciences of the United States of America*. 64:1087-1094.
- Hung, R.J., J.D. McKay, V. Gaborieau, P. Boffetta, M. Hashibe, D. Zaridze, A. Mukeria, N. Szeszenia-Dabrowska, J. Lissowska, P. Rudnai, E. Fabianova, D. Mates, V. Bencko, L. Foretova, V. Janout, C. Chen, G. Goodman, J.K. Field, T. Liloglou, G. Xinarianos, A. Cassidy, J. McLaughlin, G. Liu, S. Narod, H.E. Krokan, F. Skorpen, M.B. Elvestad, K. Hveem, L. Vatten, J. Linseisen, F. Clavel-Chapelon, P. Vineis, H.B. Bueno-de-Mesquita, E. Lund, C. Martinez, S. Bingham, T. Rasmuson, P. Hainaut, E. Riboli, W. Ahrens, S. Benhamou, P. Lagiou, D. Trichopoulos, I. Holcatova, F. Merletti, K. Kjaerheim, A. Agudo, G. Macfarlane, R. Talamini, L. Simonato, R. Lowry, D.I. Conway, A. Znaor, C. Healy, D. Zelenika, A. Boland, M. Delepine, M. Foglio, D. Lechner, F. Matsuda, H. Blanche, I. Gut, S. Heath, M. Lathrop, and P. Brennan. 2008. A susceptibility locus for lung cancer maps to nicotinic acetylcholine receptor subunit genes on 15q25. *Nature*. 452:633-637.
- Husgafvel-Pursiainen, K., P. Hackman, M. Ridanpaa, S. Anttila, A. Karjalainen, T. Partanen, O. Taikina-Aho, L. Heikkila, and H. Vainio. 1993. K-ras mutations in human adenocarcinoma of the lung: association with smoking and occupational exposure to asbestos. *International journal of cancer. Journal international du cancer*. 53:250-256.
- Hutchins, A.S., A.C. Mullen, H.W. Lee, K.J. Sykes, F.A. High, B.D. Hendrich, A.P. Bird, and S.L. Reiner. 2002. Gene silencing quantitatively controls the function of a developmental trans-activator. *Molecular cell*. 10:81-91.
- Hwang, K.K., J.C. Eissenberg, and H.J. Worman. 2001. Transcriptional repression of euchromatic genes by Drosophila heterochromatin protein 1 and histone modifiers. *Proceedings of the National Academy of Sciences of the United States of America*. 98:11423-11427.
- Ignacak, A., M. Kasztelnik, T. Sliwa, R.A. Korbut, K. Rajda, and T.J. Guzik. 2012. Prolactin--not only lactotrophin. A "new" view of the "old" hormone. *Journal of physiology and pharmacology : an official journal of the Polish Physiological Society*. 63:435-443.
- Ihle, J.N. 1996. STATs: signal transducers and activators of transcription. *Cell*. 84:331-334.
- Ihle, N.T., L.A. Byers, E.S. Kim, P. Saintigny, J.J. Lee, G.R. Blumenschein, A. Tsao, S. Liu, J.E. Larsen, J. Wang, L. Diao, K.R. Coombes, L. Chen, S. Zhang, M.F. Abdelmelek, X. Tang, V. Papadimitrakopoulou, J.D. Minna, S.M. Lippman, W.K. Hong, R.S. Herbst, Wistuba, II, J.V. Heymach, and G. Powis. 2012. Effect of KRAS oncogene substitutions on protein behavior: implications for signaling and clinical outcome. *Journal of the National Cancer Institute*. 104:228-239.
- Imbernon, E. 2003. Estimation du nombre de cas de certains cancers attribuables à des facteurs professionnels en France. Institut de Veille Sanitaire.
- INCa. 2010. La situation du cancer en France en 2010. In Rapports et synthèse - Collection états des lieux et des connaissances. Vol. Décembre 2010. Rapports et synthèse - Collection états des lieux et des connaissances
- INCa. 2011. Cancer du poumon, Bilan initial. *Collection Recommandations et référentiels*.
- INCa. 2012. La situation du cancer en France en 2012. In Rapports et synthèse - Collection états des lieux et des connaissances. Vol. Décembre 2012. Rapports et synthèse - Collection états des lieux et des connaissances.
- Ingram, D.M., E.M. Nottage, and A.N. Roberts. 1990. Prolactin and breast cancer risk. *The Medical journal of Australia*. 153:469-473.
- Ishiyama, H., A. Shibata, K. Niino, and T. Hosoya. 2004. Relationship between morphine and radiotherapy for management of symptomatic bone metastases from lung cancer. *Supportive care in cancer : official journal of the Multinational Association of Supportive Care in Cancer*. 12:743-745.

- Italiano, A., R. Attias, A. Aurias, G. Perot, F. Burel-Vandenbos, J. Otto, N. Venissac, and F. Pedetour. 2006. Molecular cytogenetic characterization of a metastatic lung sarcomatoid carcinoma: 9p23 neocentromere and 9p23-p24 amplification including JAK2 and JMJD2C. *Cancer genetics and cytogenetics*. 167:122-130.
- Jackson-Grusby, L., and R. Jaenisch. 1996. Experimental manipulation of genomic methylation. *Seminars in cancer biology*. 7:261-268.
- Jacobson, R.H., A.G. Ladurner, D.S. King, and R. Tjian. 2000. Structure and function of a human TAFII250 double bromodomain module. *Science*. 288:1422-1425.
- James, L.I., and S.V. Frye. 2013. Targeting chromatin readers. *Clinical pharmacology and therapeutics*. 93:312-314.
- James, S.R., P.A. Link, and A.R. Karpf. 2006. Epigenetic regulation of X-linked cancer/germline antigen genes by DNMT1 and DNMT3b. *Oncogene*. 25:6975-6985.
- Jang, J.S., Y.Y. Choi, W.K. Lee, J.E. Choi, S.I. Cha, Y.J. Kim, C.H. Kim, S. Kam, T.H. Jung, and J.Y. Park. 2008. Telomere length and the risk of lung cancer. *Cancer science*. 99:1385-1389.
- Jang, S.J., J.C. Soria, L. Wang, K.A. Hassan, R.C. Morice, G.L. Walsh, W.K. Hong, and L. Mao. 2001. Activation of melanoma antigen tumor antigens occurs early in lung carcinogenesis. *Cancer research*. 61:7959-7963.
- Janku, F., D.J. Stewart, and R. Kurzrock. 2010. Targeted therapy in non-small-cell lung cancer--is it becoming a reality? *Nature reviews. Clinical oncology*. 7:401-414.
- Jenuwein, T., and C.D. Allis. 2001. Translating the histone code. *Science*. 293:1074-1080.
- Jeon, H.S., Y.Y. Choi, J.E. Choi, W.K. Lee, E. Lee, S.S. Yoo, S.Y. Lee, J. Lee, S.I. Cha, C.H. Kim, and J.Y. Park. 2012. Telomere length of tumor tissues and survival in patients with early stage non-small cell lung cancer. *Molecular carcinogenesis*.
- Ji, L., B. Fang, N. Yen, K. Fong, J.D. Minna, and J.A. Roth. 1999. Induction of apoptosis and inhibition of tumorigenicity and tumor growth by adenovirus vector-mediated fragile histidine triad (FHIT) gene overexpression. *Cancer research*. 59:3333-3339.
- Jones, M.H., M. Numata, and M. Shimane. 1997. Identification and characterization of BRDT: A testis-specific gene related to the bromodomain genes RING3 and Drosophila fsh. *Genomics*. 45:529-534.
- Jones, P.A., and S.B. Baylin. 2007. The epigenomics of cancer. *Cell*. 128:683-692.
- Jones, P.A., and S.M. Taylor. 1980. Cellular differentiation, cytidine analogs and DNA methylation. *Cell*. 20:85-93.
- Juergens, R.A., J. Wrangle, F.P. Vendetti, S.C. Murphy, M. Zhao, B. Coleman, R. Sebree, K. Rodgers, C.M. Hooker, N. Franco, B. Lee, S. Tsai, I.E. Delgado, M.A. Rudek, S.A. Belinsky, J.G. Herman, S.B. Baylin, M.V. Brock, and C.M. Rudin. 2011. Combination epigenetic therapy has efficacy in patients with refractory advanced non-small cell lung cancer. *Cancer discovery*. 1:598-607.
- Jung, K.H., J.H. Noh, J.K. Kim, J.W. Eun, H.J. Bae, H.J. Xie, Y.G. Chang, M.G. Kim, H. Park, J.Y. Lee, and S.W. Nam. 2012. HDAC2 overexpression confers oncogenic potential to human lung cancer cells by deregulating expression of apoptosis and cell cycle proteins. *Journal of cellular biochemistry*. 113:2167-2177.
- Kang, Y., J.A. Hong, G.A. Chen, D.M. Nguyen, and D.S. Schrupp. 2007. Dynamic transcriptional regulatory complexes including BORIS, CTCF and Sp1 modulate NY-ESO-1 expression in lung cancer cells. *Oncogene*. 26:4394-4403.
- Kaplan, E.L.M., P. 1958. Nonparametric estimation from incomplete observations. *J. Amer. Statist. Assn*. 53:457-481.
- Kaye, F.J. 2002. RB and cyclin dependent kinase pathways: defining a distinction between RB and p16 loss in lung cancer. *Oncogene*. 21:6908-6914.
- Keeler, C., P.S. Dannies, and M.E. Hodsdon. 2003. The tertiary structure and backbone dynamics of human prolactin. *Journal of molecular biology*. 328:1105-1121.
- Kelly, P.A., J. Djiane, M.C. Postel-Vinay, and M. Ederly. 1991. The prolactin/growth hormone receptor family. *Endocrine reviews*. 12:235-251.
- Keohavong, P., M.A. DeMichele, A.C. Melacrinis, R.J. Landreneau, R.J. Weyant, and J.M. Siegfried. 1996. Detection of K-ras mutations in lung carcinomas: relationship to prognosis. *Clinical cancer research : an official journal of the American Association for Cancer Research*. 2:411-418.
- Kholmanskikh, O., A. Lorient, F. Brasseur, E. De Plaen, and C. De Smet. 2008. Expression of BORIS in melanoma: lack of association with MAGE-A1 activation. *International journal of cancer. Journal international du cancer*. 122:777-784.
- Khorasanizadeh, S. 2004. The nucleosome: from genomic organization to genomic regulation. *Cell*. 116:259-272.
- Kiefer, J.C. 2007. Back to basics: Sox genes. *Developmental dynamics : an official publication of the American Association of Anatomists*. 236:2356-2366.
- Kikuchi, R., S. Yagi, H. Kusuhara, S. Imai, Y. Sugiyama, and K. Shiota. 2010. Genome-wide analysis of epigenetic signatures for kidney-specific transporters. *Kidney international*. 78:569-577.
- Kim, C.F., E.L. Jackson, A.E. Woolfenden, S. Lawrence, I. Babar, S. Vogel, D. Crowley, R.T. Bronson, and T. Jacks. 2005. Identification of bronchioalveolar stem cells in normal lung and lung cancer. *Cell*. 121:823-835.
- Kim, E., and W. Deppert. 2006. The versatile interactions of p53 with DNA: when flexibility serves specificity. *Cell death and differentiation*. 13:885-889.
- Kim, H., Y.M. Kwon, J.S. Kim, J. Han, Y.M. Shim, J. Park, and D.H. Kim. 2006a. Elevated mRNA levels of DNA methyltransferase-1 as an independent prognostic factor in primary nonsmall cell lung cancer. *Cancer*. 107:1042-1049.
- Kim, H.R., E.J. Kim, S.H. Yang, E.T. Jeong, C. Park, J.H. Lee, M.J. Youn, H.S. So, and R. Park. 2006b. Trichostatin A induces apoptosis in lung cancer cells via simultaneous activation of the death receptor-mediated and mitochondrial pathway? *Experimental & molecular medicine*. 38:616-624.

- Kimura, K., A. Wakamatsu, Y. Suzuki, T. Ota, T. Nishikawa, R. Yamashita, J. Yamamoto, M. Sekine, K. Tsuritani, H. Wakaguri, S. Ishii, T. Sugiyama, K. Saito, Y. Isono, R. Irie, N. Kushida, T. Yoneyama, R. Otsuka, K. Kanda, T. Yokoi, H. Kondo, M. Wagatsuma, K. Murakawa, S. Ishida, T. Ishibashi, A. Takahashi-Fujii, T. Tanase, K. Nagai, H. Kikuchi, K. Nakai, T. Isogai, and S. Sugano. 2006. Diversification of transcriptional modulation: large-scale identification and characterization of putative alternative promoters of human genes. *Genome research*. 16:55-65.
- Kishimoto, M., T. Kohno, K. Okudela, A. Otsuka, H. Sasaki, C. Tanabe, T. Sakiyama, C. Hirama, I. Kitabayashi, J.D. Minna, S. Takenoshita, and J. Yokota. 2005. Mutations and deletions of the CBP gene in human lung cancer. *Clinical cancer research : an official journal of the American Association for Cancer Research*. 11:512-519.
- Kitamura, H., T. Yazawa, K. Okudela, H. Shimoyamada, and H. Sato. 2008. Molecular and Genetic Pathogenesis of Lung Cancer: Differences Between Small-Cell and Non-Small-Cell Carcinomas. *The Open Pathology Journal*.
- Klose, R.J., and A.P. Bird. 2006. Genomic DNA methylation: the mark and its mediators. *Trends in biochemical sciences*. 31:89-97.
- Koenig, K.L., P. Toniolo, P.F. Bruning, J.M. Bonfrer, R.E. Shore, and B.S. Pasternack. 1993. Reliability of serum prolactin measurements in women. *Cancer epidemiology, biomarkers & prevention : a publication of the American Association for Cancer Research, cosponsored by the American Society of Preventive Oncology*. 2:411-414.
- Kohno, T., A. Otsuka, L. Girard, M. Sato, R. Iwakawa, H. Ogiwara, M. Sanchez-Cespedes, J.D. Minna, and J. Yokota. 2010. A catalog of genes homozygously deleted in human lung cancer and the candidacy of PTPRD as a tumor suppressor gene. *Genes, chromosomes & cancer*. 49:342-352.
- Kohno, T., and J. Yokota. 1999. How many tumor suppressor genes are involved in human lung carcinogenesis? *Carcinogenesis*. 20:1403-1410.
- Kok, K., S.L. Naylor, and C.H. Buys. 1997. Deletions of the short arm of chromosome 3 in solid tumors and the search for suppressor genes. *Advances in cancer research*. 71:27-92.
- Koop, R., L. Di Croce, and M. Beato. 2003. Histone H1 enhances synergistic activation of the MMTV promoter in chromatin. *The EMBO journal*. 22:588-599.
- Kornberg, R.D., and J.O. Thomas. 1974. Chromatin structure; oligomers of the histones. *Science*. 184:865-868.
- Kouzarides, T. 2007a. Chromatin modifications and their function. *Cell*. 128:693-705.
- Kouzarides, T. 2007b. SnapShot: Histone-modifying enzymes. *Cell*. 131:822.
- Kuehner, J.N., E.L. Pearson, and C. Moore. 2011. Unravelling the means to an end: RNA polymerase II transcription termination. *Nature reviews. Molecular cell biology*. 12:283-294.
- Kwak, E.L.D.R.C., J. Clark, G. I. Shapiro, R. G. Maki, M. J. Ratain, B. Solomon, Y. Bang, S. Ou, R. Salgia. 2009. Clinical activity observed in a phase I dose escalation trial of an oral c-met and ALK inhibitor, PF-02341066. *J Clin Oncol* 27:15s, 2009 (suppl; abstr 3509). *J Clin Oncol* 27:15s, 2009 (suppl; abstr 3509).
- L. Leonart, M., F. Vidal, D. Gallardo, M. Diaz-Fuertes, F. Rojo, M. Cuatrecasas, L. Lopez-Vicente, H. Kondoh, C. Blanco, A. Carnero, and S. Ramon y Cajal. 2006. New p53 related genes in human tumors: significant downregulation in colon and lung carcinomas. *Oncology reports*. 16:603-608.
- La Salle, S., K. Palmer, M. O'Brien, J.C. Schimenti, J. Eppig, and M.A. Handel. 2012. Spata22, a novel vertebrate-specific gene, is required for meiotic progress in mouse germ cells. *Biology of reproduction*. 86:45.
- Landry, J.R., D.L. Mager, and B.T. Wilhelm. 2003. Complex controls: the role of alternative promoters in mammalian genomes. *Trends in genetics : TIG*. 19:640-648.
- Lantuejoul, S., C. Salon, J.C. Soria, and E. Brambilla. 2007. Telomerase expression in lung preneoplasia and neoplasia. *International journal of cancer. Journal international du cancer*. 120:1835-1841.
- Lantuejoul, S., J.C. Soria, D. Moro-Sibilot, L. Morat, S. Veyrenc, P. Lorimier, P.Y. Brichon, L. Sabatier, C. Brambilla, and E. Brambilla. 2004. Differential expression of telomerase reverse transcriptase (hTERT) in lung tumours. *British journal of cancer*. 90:1222-1229.
- Lantu  joul, S.B., E. 2006. Prognostic biomarkers in non-small-cell lung carcinoma. *Current Diagnostic Pathology*. Volume 12:p. 418-428.
- Latham, J.A., and S.Y. Dent. 2007. Cross-regulation of histone modifications. *Nature structural & molecular biology*. 14:1017-1024.
- Latil, A., D. Vidaud, A. Valeri, G. Fournier, M. Vidaud, R. Lidereau, O. Cussenot, and I. Biache. 2000. htert expression correlates with MYC over-expression in human prostate cancer. *International journal of cancer. Journal international du cancer*. 89:172-176.
- Lee, J.H., M.L. Choy, and P.A. Marks. 2012. Mechanisms of resistance to histone deacetylase inhibitors. *Advances in cancer research*. 116:39-86.
- Lee, S.H., J. Kunz, S.H. Lin, and L.Y. Yu-Lee. 2007. 16-kDa prolactin inhibits endothelial cell migration by down-regulating the Ras-Tiam1-Rac1-Pak1 signaling pathway. *Cancer research*. 67:11045-11053.
- Lee, T.I., and R.A. Young. 1998. Regulation of gene expression by TBP-associated proteins. *Genes & development*. 12:1398-1408.
- Lee, W., Z. Jiang, J. Liu, P.M. Haverty, Y. Guan, J. Stinson, P. Yue, Y. Zhang, K.P. Pant, D. Bhatt, C. Ha, S. Johnson, M.I. Kennemer, S. Mohan, I. Nazarenko, C. Watanabe, A.B. Sparks, D.S. Shames, R. Gentleman, F.J. de Sauvage, H. Stern, A. Pandita, D.G. Ballinger, R. Drmanac, Z. Modrusan, S. Seshagiri, and Z. Zhang. 2010. The mutation spectrum revealed by paired genome sequences from a lung cancer patient. *Nature*. 465:473-477.

- Lehnertz, B., Y. Ueda, A.A. Derijck, U. Braunschweig, L. Perez-Burgos, S. Kubicek, T. Chen, E. Li, T. Jenuwein, and A.H. Peters. 2003. Suv39h-mediated histone H3 lysine 9 methylation directs DNA methylation to major satellite repeats at pericentric heterochromatin. *Current biology : CB*. 13:1192-1200.
- Lengauer, C., K.W. Kinzler, and B. Vogelstein. 1998. Genetic instabilities in human cancers. *Nature*. 396:643-649.
- Levin, P.A., and W.B. Malarkey. 1981. Daughters of women with breast cancer have elevated mean 24-hour prolactin (PRL) levels and a partial resistance of PRL to dopamine suppression. *The Journal of clinical endocrinology and metabolism*. 53:179-183.
- Levina, V.V., B. Nolen, Y. Su, A.K. Godwin, D. Fishman, J. Liu, G. Mor, L.G. Maxwell, R.B. Herberman, M.J. Szczepanski, M.E. Szajnik, E. Gorelik, and A.E. Lokshin. 2009. Biological significance of prolactin in gynecologic cancers. *Cancer research*. 69:5226-5233.
- Levine, M., and R. Tjian. 2003. Transcription regulation and animal diversity. *Nature*. 424:147-151.
- Lewis, A., K. Mitsuya, M. Constanica, and W. Reik. 2004. Tandem repeat hypothesis in imprinting: deletion of a conserved direct repeat element upstream of H19 has no effect on imprinting in the Igf2-H19 region. *Molecular and cellular biology*. 24:5650-5656.
- Li, F., B. Ye, L. Hong, H. Xu, and M.C. Fishbein. 2011. Epigenetic modifications of histone h4 in lung neuroendocrine tumors. *Applied immunohistochemistry & molecular morphology : AIMM / official publication of the Society for Applied Immunohistochemistry*. 19:389-394.
- Lim, E.H., S.L. Zhang, J.L. Li, W.S. Yap, T.C. Howe, B.P. Tan, Y.S. Lee, D. Wong, K.L. Khoo, K.Y. Seto, L. Tan, T. Agasthian, H.N. Koong, J. Tam, C. Tan, M. Caleb, A. Chang, A. Ng, and P. Tan. 2009. Using whole genome amplification (WGA) of low-volume biopsies to assess the prognostic role of EGFR, KRAS, p53, and CMET mutations in advanced-stage non-small cell lung cancer (NSCLC). *Journal of thoracic oncology : official publication of the International Association for the Study of Lung Cancer*. 4:12-21.
- Lin, J., M. Lai, Q. Huang, Y. Ma, J. Cui, and W. Ruan. 2007a. Methylation patterns of IGFBP7 in colon cancer cell lines are associated with levels of gene expression. *The Journal of pathology*. 212:83-90.
- Lin, R.K., H.S. Hsu, J.W. Chang, C.Y. Chen, J.T. Chen, and Y.C. Wang. 2007b. Alteration of DNA methyltransferases contributes to 5'CpG methylation and poor prognosis in lung cancer. *Lung cancer*. 55:205-213.
- Lin, T.S., H. Lee, R.A. Chen, M.L. Ho, C.Y. Lin, Y.H. Chen, Y.Y. Tsai, M.C. Chou, and Y.W. Cheng. 2005. An association of DNMT3b protein expression with P16INK4a promoter hypermethylation in non-smoking female lung cancer with human papillomavirus infection. *Cancer letters*. 226:77-84.
- Linnoila, R.I. 1996. Spectrum of neuroendocrine differentiation in lung cancer cell lines featured by cytomorphology, markers, and their corresponding tumors. *Journal of cellular biochemistry. Supplement*. 24:92-106.
- Liu, F., J.K. Killian, M. Yang, R.L. Walker, J.A. Hong, M. Zhang, S. Davis, Y. Zhang, M. Hussain, S. Xi, M. Rao, P.A. Meltzer, and D.S. Schrupp. 2010. Epigenomic alterations and gene expression profiles in respiratory epithelia exposed to cigarette smoke condensate. *Oncogene*. 29:3650-3664.
- Lkhider, M., T. Seddiki, and M. Ollivier-Bousquet. 2010. [Prolactin and its cleaved 16 kDa fragment]. *Medecine sciences : M/S*. 26:1049-1055.
- Lodish, H.F., Berk A., Kaiser C., Krieger M., Scott M., Bretscher A., Ploegh H., Matsudaira P. 2008. Molecular Cell Biology. W. H.. Freeman and Company, N.Y.
- Loeb, L.A. 1991. Mutator phenotype may be required for multistage carcinogenesis. *Cancer research*. 51:3075-3079.
- Loeb, L.A. 2011. Human cancers express mutator phenotypes: origin, consequences and targeting. *Nature reviews. Cancer*. 11:450-457.
- Lorincz, M.C., D.R. Dickerson, M. Schmitt, and M. Groudine. 2004. Intragenic DNA methylation alters chromatin structure and elongation efficiency in mammalian cells. *Nature structural & molecular biology*. 11:1068-1075.
- Loriot, A., E. De Plaen, T. Boon, and C. De Smet. 2006. Transient down-regulation of DNMT1 methyltransferase leads to activation and stable hypomethylation of MAGE-A1 in melanoma cells. *The Journal of biological chemistry*. 281:10118-10126.
- Lou-Qian, Z., Y. Rong, L. Ming, Y. Xin, J. Feng, and X. Lin. 2013. The prognostic value of epigenetic silencing of p16 gene in NSCLC patients: a systematic review and meta-analysis. *PloS one*. 8:e54970.
- Love, R.R., and D.P. Rose. 1985. Elevated bioactive prolactin in women at risk for familial breast cancer. *European journal of cancer & clinical oncology*. 21:1553-1554.
- Lu, Y., S. Cheng, and S. Guo. 1996. [C-myc activation by translocation in human primary non-small cell lung cancer]. *Zhonghua yi xue za zhi*. 76:275-277.
- Luch, A., and W.M. Baird. 2005. Metabolic activation and detoxification of polycyclic aromatic hydrocarbons. In *The carcinogenic effects of polycyclic aromatic hydrocarbons*. L.I.C. Press, editor. 19-96.
- Luger, K., A.W. Mader, R.K. Richmond, D.F. Sargent, and T.J. Richmond. 1997. Crystal structure of the nucleosome core particle at 2.8 Å resolution. *Nature*. 389:251-260.
- Lyons, D.J., A. Hellysaz, and C. Broberger. 2012. Prolactin regulates tuberoinfundibular dopamine neuron discharge pattern: novel feedback control mechanisms in the lactotrophic axis. *The Journal of neuroscience : the official journal of the Society for Neuroscience*. 32:8074-8083.
- Mabry, M., B.D. Nelkin, J.P. Falco, L.F. Barr, and S.B. Baylin. 1991. Transitions between lung cancer phenotypes--implications for tumor progression. *Cancer cells*. 3:53-58.
- MacCallum, C., and H.H. Gillenwater. 2006. Second-line treatment of small-cell lung cancer. *Current oncology reports*. 8:258-264.

- Macotella, Y., M.B. Aguilar, J. Guzman-Morales, J.C. Rivera, C. Zermeno, F. Lopez-Barrera, G. Nava, C. Lavalle, G. Martinez de la Escalera, and C. Clapp. 2006. Matrix metalloproteases from chondrocytes generate an antiangiogenic 16 kDa prolactin. *Journal of cell science*. 119:1790-1800.
- Maison, C., and G. Almouzni. 2004. HP1 and the dynamics of heterochromatin maintenance. *Nature reviews. Molecular cell biology*. 5:296-304.
- Malik, S., and R.G. Roeder. 2010. The metazoan Mediator co-activator complex as an integrative hub for transcriptional regulation. *Nature reviews. Genetics*. 11:761-772.
- Mancini, T., F.F. Casanueva, and A. Giustina. 2008. Hyperprolactinemia and prolactinomas. *Endocrinology and metabolism clinics of North America*. 37:67-99, viii.
- Mao, L., J.S. Lee, J.M. Kurie, Y.H. Fan, S.M. Lippman, J.J. Lee, J.Y. Ro, A. Broxson, R. Yu, R.C. Morice, B.L. Kemp, F.R. Khuri, G.L. Walsh, W.N. Hittelman, and W.K. Hong. 1997. Clonal genetic alterations in the lungs of current and former smokers. *Journal of the National Cancer Institute*. 89:857-862.
- Marchand, M., N. van Baren, P. Weynants, V. Brichard, B. Dreno, M.H. Tessier, E. Rankin, G. Parmiani, F. Arienti, Y. Humblet, A. Bourlond, R. Vanwijck, D. Lienard, M. Beauduin, P.Y. Dietrich, V. Russo, J. Kerger, G. Masucci, E. Jager, J. De Greve, J. Atzpodien, F. Brasseur, P.G. Coulie, P. van der Bruggen, and T. Boon. 1999. Tumor regressions observed in patients with metastatic melanoma treated with an antigenic peptide encoded by gene MAGE-3 and presented by HLA-A1. *International journal of cancer. Journal international du cancer*. 80:219-230.
- Marks, P.A., and R. Breslow. 2007. Dimethyl sulfoxide to vorinostat: development of this histone deacetylase inhibitor as an anticancer drug. *Nature biotechnology*. 25:84-90.
- Marks, P.A., V.M. Richon, and R.A. Rifkind. 2000. Histone deacetylase inhibitors: inducers of differentiation or apoptosis of transformed cells. *Journal of the National Cancer Institute*. 92:1210-1216.
- Martens, J.A., and F. Winston. 2003. Recent advances in understanding chromatin remodeling by Swi/Snf complexes. *Current opinion in genetics & development*. 13:136-142.
- Martin, G.S. 1970. Rous sarcoma virus: a function required for the maintenance of the transformed state. *Nature*. 227:1021-1023.
- Mascaux, C., N. Peled, K. Garg, Y. Kato, M.W. Wynes, and F.R. Hirsch. 2010. Early detection and screening of lung cancer. *Expert review of molecular diagnostics*. 10:799-815.
- Masuda, A., and T. Takahashi. 2002. Chromosome instability in human lung cancers: possible underlying mechanisms and potential consequences in the pathogenesis. *Oncogene*. 21:6884-6897.
- Maurer, P., F. T'Sas, L. Coutte, N. Callens, C. Brenner, C. Van Lint, Y. de Launoit, and J.L. Baert. 2003. FEV acts as a transcriptional repressor through its DNA-binding ETS domain and alanine-rich domain. *Oncogene*. 22:3319-3329.
- Mavilio, F., A. Giampaolo, A. Care, G. Migliaccio, M. Calandrini, G. Russo, G.L. Pagliardi, G. Mastroberardino, M. Marinucci, and C. Peschle. 1983. Molecular mechanisms of human hemoglobin switching: selective undermethylation and expression of globin genes in embryonic, fetal, and adult erythroblasts. *Proceedings of the National Academy of Sciences of the United States of America*. 80:6907-6911.
- McDermott, U., A.J. Iafrate, N.S. Gray, T. Shioda, M. Classon, S. Maheswaran, W. Zhou, H.G. Choi, S.L. Smith, L. Dowell, L.E. Ulkus, G. Kuhlmann, P. Greninger, J.G. Christensen, D.A. Haber, and J. Settleman. 2008. Genomic alterations of anaplastic lymphoma kinase may sensitize tumors to anaplastic lymphoma kinase inhibitors. *Cancer research*. 68:3389-3395.
- Melloni, G., A.J. Ferreri, V. Russo, L. Gattinoni, G. Arrigoni, G.L. Ceresoli, P. Zannini, and C. Traversari. 2004. Prognostic significance of cancer-testis gene expression in resected non-small cell lung cancer patients. *Oncology reports*. 12:145-151.
- Menegoz, F., and L. Chérié-Challine. 1998. Le cancer en France: incidence et mortalité. M.d.l.e.e.d.l. solidarité, editor.
- Messini, C.I., K. Dafopoulos, N. Chalvatzas, P. Georgoulas, G. Anifandis, and I.E. Messinis. 2010. Effect of ghrelin and thyrotropin-releasing hormone on prolactin secretion in normal women. *Hormone and metabolic research = Hormon- und Stoffwechselforschung = Hormones et métabolisme*. 42:204-208.
- Miki, K., E. Shimizu, S. Yano, K. Tani, and S. Sone. 2001. Demethylation by 5-aza-2'-deoxycytidine (5-azadC) of p16INK4A gene results in downregulation of vascular endothelial growth factor expression in human lung cancer cell lines. *Oncology research*. 12:335-342.
- Minna, J.D., J.A. Roth, and A.F. Gazdar. 2002. Focus on lung cancer. *Cancer cell*. 1:49-52.
- Minna, J.D., and J.H. Schiller. 2008. Harrison's Principles of Internal Medicine.
- Mino-Kenudson, M., L.R. Chirieac, K. Law, J.L. Hornick, N. Lindeman, E.J. Mark, D.W. Cohen, B.E. Johnson, P.A. Janne, A.J. Iafrate, and S.J. Rodig. 2010. A novel, highly sensitive antibody allows for the routine detection of ALK-rearranged lung adenocarcinomas by standard immunohistochemistry. *Clinical cancer research : an official journal of the American Association for Cancer Research*. 16:1561-1571.
- Momparler, R.L., N. Eliopoulos, and J. Ayoub. 2000. Evaluation of an inhibitor of DNA methylation, 5-aza-2'-deoxycytidine, for the treatment of lung cancer and the future role of gene therapy. *Advances in experimental medicine and biology*. 465:433-446.
- Monte, M., M. Simonatto, L.Y. Peche, D.R. Bublik, S. Gobessi, M.A. Pierotti, M. Rodolfo, and C. Schneider. 2006. MAGE-A tumor antigens target p53 transactivation function through histone deacetylase recruitment and confer resistance to chemotherapeutic agents. *Proceedings of the National Academy of Sciences of the United States of America*. 103:11160-11165.

- Montellier, E., S. Rousseaux, Y. Zhao, and S. Khochbin. 2012. Histone crotonylation specifically marks the haploid male germ cell gene expression program: post-meiotic male-specific gene expression. *BioEssays : news and reviews in molecular, cellular and developmental biology*. 34:187-193.
- Montellier, E.B., F.; Rousseaux, S.; Zhang, K.; Buchou, T.; Fenaille, F.; Shiota, H.; Debernardi, A.; Héry, P.; Holota, H.; Bergon, A.; Lopez, F.; Guardiola, P.; Pernet, K.; Imbert, J.; Petosa, C.; Tan, M.; Zhao, Y.; Gérard, M.; Khochbin, S. 2013. Chromatin-to-nucleoprotamine transition is controlled by the histone H2B variant TH2B. *Submitted*.
- Morgunkova, A., and N.A. Barlev. 2006. Lysine methylation goes global. *Cell cycle*. 5:1308-1312.
- Moriniere, J., S. Rousseaux, U. Steuerwald, M. Soler-Lopez, S. Curtet, A.L. Vitte, J. Govin, J. Gaucher, K. Sadoul, D.J. Hart, J. Krijgsveld, S. Khochbin, C.W. Muller, and C. Petosa. 2009. Cooperative binding of two acetylation marks on a histone tail by a single bromodomain. *Nature*. 461:664-668.
- Morison, I.M., J.P. Ramsay, and H.G. Spencer. 2005. A census of mammalian imprinting. *Trends in genetics : TIG*. 21:457-465.
- Morris, S.W., C. Naeve, P. Mathew, P.L. James, M.N. Kirstein, X. Cui, and D.P. Witte. 1997. ALK, the chromosome 2 gene locus altered by the t(2;5) in non-Hodgkin's lymphoma, encodes a novel neural receptor tyrosine kinase that is highly related to leukocyte tyrosine kinase (LTK). *Oncogene*. 14:2175-2188.
- Motadi, L.R., N.L. Misso, Z. Dlamini, and K.D. Bhoola. 2007. Molecular genetics and mechanisms of apoptosis in carcinomas of the lung and pleura: therapeutic targets. *International immunopharmacology*. 7:1934-1947.
- Mountain, C., H. Libshitz, and E. Hermes. 2003. A Handbook for Staging, Imaging, and Lymph Node Classification.
- Mukhopadhyay, N.K., E. Weisberg, D. Gilchrist, R. Bueno, D.J. Sugarbaker, and M.T. Jaklitsch. 2006. Effectiveness of trichostatin A as a potential candidate for anticancer therapy in non-small-cell lung cancer. *The Annals of thoracic surgery*. 81:1034-1042.
- Murray, N., and A.T. Turrisi, 3rd. 2006. A review of first-line treatment for small-cell lung cancer. *Journal of thoracic oncology : official publication of the International Association for the Study of Lung Cancer*. 1:270-278.
- Narlikar, G.J., H.Y. Fan, and R.E. Kingston. 2002. Cooperation between complexes that regulate chromatin structure and transcription. *Cell*. 108:475-487.
- Naylor, S.L., B.E. Johnson, J.D. Minna, and A.Y. Sakaguchi. 1987. Loss of heterozygosity of chromosome 3p markers in small-cell lung cancer. *Nature*. 329:451-454.
- Neilson, L.M., J. Zhu, J. Xie, M.G. Malabarba, K. Sakamoto, K.U. Wagner, R.A. Kirken, and H. Rui. 2007. Coactivation of janus tyrosine kinase (Jak)1 positively modulates prolactin-Jak2 signaling in breast cancer: recruitment of ERK and signal transducer and activator of transcription (Stat)3 and enhancement of Akt and Stat5a/b pathways. *Molecular endocrinology*. 21:2218-2232.
- Nelson, H.H., J.K. Wiencke, L. Gunn, J.C. Wain, D.C. Christiani, and K.T. Kelsey. 1998. Chromosome 3p14 alterations in lung cancer: evidence that FHIT exon deletion is a target of tobacco carcinogens and asbestos. *Cancer research*. 58:1804-1807.
- Nevalainen, M.T., E.M. Valve, P.M. Ingleton, M. Nurmi, P.M. Martikainen, and P.L. Harkonen. 1997. Prolactin and prolactin receptors are expressed and functioning in human prostate. *The Journal of clinical investigation*. 99:618-627.
- Ng, H.H., Y. Zhang, B. Hendrich, C.A. Johnson, B.M. Turner, H. Erdjument-Bromage, P. Tempst, D. Reinberg, and A. Bird. 1999. MBD2 is a transcriptional repressor belonging to the MeCP1 histone deacetylase complex. *Nature genetics*. 23:58-61.
- Niall, H.D., M.L. Hogan, R. Sauer, I.Y. Rosenblum, and F.C. Greenwood. 1971. Sequences of pituitary and placental lactogenic and growth hormones: evolution from a primordial peptide by gene reduplication. *Proceedings of the National Academy of Sciences of the United States of America*. 68:866-870.
- Nielsen, A.L., J.A. Ortiz, J. You, M. Oulad-Abdelghani, R. Khechumian, A. Gansmuller, P. Chambon, and R. Losson. 1999. Interaction with members of the heterochromatin protein 1 (HP1) family and histone deacetylation are differentially involved in transcriptional silencing by members of the TIF1 family. *The EMBO journal*. 18:6385-6395.
- Nieto, M.A. 2002. The snail superfamily of zinc-finger transcription factors. *Nature reviews. Molecular cell biology*. 3:155-166.
- Noma, K., C.D. Allis, and S.I. Grewal. 2001. Transitions in distinct histone H3 methylation patterns at the heterochromatin domain boundaries. *Science*. 293:1150-1155.
- Nuber, U.A., S. Kriaucionis, T.C. Roloff, J. Guy, J. Selfridge, C. Steinhoff, R. Schulz, B. Lipkowitz, H.H. Ropers, M.C. Holmes, and A. Bird. 2005. Up-regulation of glucocorticoid-regulated genes in a mouse model of Rett syndrome. *Human molecular genetics*. 14:2247-2256.
- Nuovo, G.J., T.W. Plaia, S.A. Belinsky, S.B. Baylin, and J.G. Herman. 1999. In situ detection of the hypermethylation-induced inactivation of the p16 gene as an early event in oncogenesis. *Proceedings of the National Academy of Sciences of the United States of America*. 96:12754-12759.
- Oakes, S.R., F.G. Robertson, J.G. Kench, M. Gardiner-Garden, M.P. Wand, J.E. Green, and C.J. Ormandy. 2007. Loss of mammary epithelial prolactin receptor delays tumor formation by reducing cell proliferation in low-grade preinvasive lesions. *Oncogene*. 26:543-553.
- Ogino, S., C.S. Fuchs, and E. Giovannucci. 2012. How many molecular subtypes? Implications of the unique tumor principle in personalized medicine. *Expert review of molecular diagnostics*. 12:621-628.
- Ogryzko, V.V., R.L. Schiltz, V. Russanova, B.H. Howard, and Y. Nakatani. 1996. The transcriptional coactivators p300 and CBP are histone acetyltransferases. *Cell*. 87:953-959.

- Ohno, S., W.D. Kaplan, and R. Kinoshita. 1959. Formation of the sex chromatin by a single X-chromosome in liver cells of *Rattus norvegicus*. *Experimental cell research*. 18:415-418.
- Okamoto, A., D.J. Demetrick, E.A. Spillare, K. Hagiwara, S.P. Hussain, W.P. Bennett, K. Forrester, B. Gerwin, M. Serrano, D.H. Beach, and et al. 1994. Mutations and altered expression of p16INK4 in human cancer. *Proceedings of the National Academy of Sciences of the United States of America*. 91:11045-11049.
- Olaussen, K.A., A. Dunant, P. Fouret, E. Brambilla, F. Andre, V. Haddad, E. Taranchon, M. Filipits, R. Pirker, H.H. Popper, R. Stahel, L. Sabatier, J.P. Pignon, T. Tursz, T. Le Chevalier, J.C. Soria, and I.B. Investigators. 2006. DNA repair by ERCC1 in non-small-cell lung cancer and cisplatin-based adjuvant chemotherapy. *The New England journal of medicine*. 355:983-991.
- Old, L.J. 2001. Cancer/testis (CT) antigens - a new link between gametogenesis and cancer. *Cancer immunity*. 1:1.
- Olins, A.L., and D.E. Olins. 1974. Spheroid chromatin units (v bodies). *Science*. 183:330-332.
- Ollinger, R., M. Alsheimer, and R. Benavente. 2005. Mammalian protein SCP1 forms synaptonemal complex-like structures in the absence of meiotic chromosomes. *Molecular biology of the cell*. 16:212-217.
- Ooi, S.K., C. Qiu, E. Bernstein, K. Li, D. Jia, Z. Yang, H. Erdjument-Bromage, P. Tempst, S.P. Lin, C.D. Allis, X. Cheng, and T.H. Bestor. 2007. DNMT3L connects unmethylated lysine 4 of histone H3 to de novo methylation of DNA. *Nature*. 448:714-717.
- Osada, H., Y. Tatematsu, H. Saito, Y. Yatabe, T. Mitsudomi, and T. Takahashi. 2004. Reduced expression of class II histone deacetylase genes is associated with poor prognosis in lung cancer patients. *International journal of cancer. Journal international du cancer*. 112:26-32.
- Osaki, E., Y. Nishina, J. Inazawa, N.G. Copeland, D.J. Gilbert, N.A. Jenkins, M. Ohsugi, T. Tezuka, M. Yoshida, and K. Semba. 1999. Identification of a novel Sry-related gene and its germ cell-specific expression. *Nucleic acids research*. 27:2503-2510.
- Osterlind, K., and P.K. Andersen. 1986. Prognostic factors in small cell lung cancer: multivariate model based on 778 patients treated with chemotherapy with or without irradiation. *Cancer research*. 46:4189-4194.
- Oudet, P., M. Gross-Bellard, and P. Chambon. 1975. Electron microscopic and biochemical evidence that chromatin structure is a repeating unit. *Cell*. 4:281-300.
- Owerbach, D., W.J. Rutter, N.E. Cooke, J.A. Martial, and T.B. Shows. 1981. The prolactin gene is located on chromosome 6 in humans. *Science*. 212:815-816.
- Palangat, M., and D.R. Larson. 2012. Complexity of RNA polymerase II elongation dynamics. *Biochimica et biophysica acta*. 1819:667-672.
- Palmisano, W.A., K.K. Divine, G. Saccomanno, F.D. Gilliland, S.B. Baylin, J.G. Herman, and S.A. Belinsky. 2000. Predicting lung cancer by detecting aberrant promoter methylation in sputum. *Cancer research*. 60:5954-5958.
- Pao, W., V.A. Miller, K.A. Politi, G.J. Riely, R. Somwar, M.F. Zakowski, M.G. Kris, and H. Varmus. 2005. Acquired resistance of lung adenocarcinomas to gefitinib or erlotinib is associated with a second mutation in the EGFR kinase domain. *PLoS medicine*. 2:e73.
- Park, J.H., G.H. Kong, and S.W. Lee. 2002. hMAGE-A1 overexpression reduces TNF-alpha cytotoxicity in ME-180 cells. *Molecules and cells*. 14:122-129.
- Parker, D.C., L.G. Rossman, and E.F. Vander Laan. 1973. Sleep-related, nycthermal and briefly episodic variation in human plasma prolactin concentrations. *The Journal of clinical endocrinology and metabolism*. 36:1119-1124.
- Patel, J.D. 2005. Lung cancer in women. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology*. 23:3212-3218.
- Peinado, H., E. Ballestar, M. Esteller, and A. Cano. 2004. Snail mediates E-cadherin repression by the recruitment of the Sin3A/histone deacetylase 1 (HDAC1)/HDAC2 complex. *Molecular and cellular biology*. 24:306-319.
- Peinado, H., D. Olmeda, and A. Cano. 2007. Snail, Zeb and bHLH factors in tumour progression: an alliance against the epithelial phenotype? *Nature reviews. Cancer*. 7:415-428.
- Pellegrini, S., and I. Dusanter-Fourt. 1997. The structure, regulation and function of the Janus kinases (JAKs) and the signal transducers and activators of transcription (STATs). *European journal of biochemistry / FEBS*. 248:615-633.
- Pfaffl, M.W. 2001. A new mathematical model for relative quantification in real-time RT-PCR. *Nucleic acids research*. 29:e45.
- Pfeifer, G.P., M.F. Denissenko, M. Olivier, N. Tretyakova, S.S. Hecht, and P. Hainaut. 2002. Tobacco smoke carcinogens, DNA damage and p53 mutations in smoking-associated cancers. *Oncogene*. 21:7435-7451.
- Pfeifer, G.P., and T.A. Rauch. 2009. DNA methylation patterns in lung carcinomas. *Seminars in cancer biology*. 19:181-187.
- Pfützner, E., R. Jahne, M. Wissler, E. Stoecklin, and B. Groner. 1998. p300/CREB-binding protein enhances the prolactin-mediated transcriptional induction through direct interaction with the transactivation domain of Stat5, but does not participate in the Stat5-mediated suppression of the glucocorticoid response. *Molecular endocrinology*. 12:1582-1593.
- Picaud, S., D. Da Costa, A. Thanasopoulou, P. Filippakopoulos, P.V. Fish, M. Philpott, O. Fedorov, P. Brennan, M.E. Bunnage, D.R. Owen, J.E. Bradner, P. Taniere, B. O'Sullivan, S. Muller, J. Schwaller, T. Stankovic, and S. Knapp. 2013. PFI-1, a Highly Selective Protein Interaction Inhibitor, Targeting BET Bromodomains. *Cancer research*.
- Pilotto, S., U. Peretti, S. Novello, G. Rossi, M. Milella, M. Gaj Levra, L. Ciuffreda, F. Massari, M. Brunelli, G. Tortora, and E. Bria. 2013. PROFILING non-small-cell lung cancer patients for treatment with crizotinib according to anaplastic lymphoma kinase abnormalities: translating science into medicine. *Expert opinion on pharmacotherapy*. 14:597-608.

- Pirozynski, M. 2006. 100 years of lung cancer. *Respiratory medicine*. 100:2073-2084.
- Pivot-Pajot, C., C. Caron, J. Govin, A. Vion, S. Rousseaux, and S. Khochbin. 2003. Acetylation-dependent chromatin reorganization by BRDT, a testis-specific bromodomain-containing protein. *Molecular and cellular biology*. 23:5354-5365.
- Piwnicka, D., I. Fernandez, N. Binart, P. Touraine, P.A. Kelly, and V. Goffin. 2006. A new mechanism for prolactin processing into 16K PRL by secreted cathepsin D. *Molecular endocrinology*. 20:3263-3278.
- Pleasance, E.D., P.J. Stephens, S. O'Meara, D.J. McBride, A. Meynert, D. Jones, M.L. Lin, D. Beare, K.W. Lau, C. Greenman, I. Varela, S. Nik-Zainal, H.R. Davies, G.R. Ordóñez, L.J. Mudie, C. Latimer, S. Edkins, L. Stebbings, L. Chen, M. Jia, C. Leroy, J. Marshall, A. Menzies, A. Butler, J.W. Teague, J. Mangion, Y.A. Sun, S.F. McLaughlin, H.E. Peckham, E.F. Tsung, G.L. Costa, C.C. Lee, J.D. Minna, A. Gazdar, E. Birney, M.D. Rhodes, K.J. McKernan, M.R. Stratton, P.A. Futreal, and P.J. Campbell. 2010. A small-cell lung cancer genome with complex signatures of tobacco exposure. *Nature*. 463:184-190.
- Posner, B.I., P.A. Kelly, R.P. Shiu, and H.G. Friesen. 1974. Studies of insulin, growth hormone and prolactin binding: tissue distribution, species variation and characterization. *Endocrinology*. 95:521-531.
- Postel-Vinay, S., E. Vanhecke, K.A. Olaussen, C.J. Lord, A. Ashworth, and J.C. Soria. 2012. The potential of exploiting DNA-repair defects for optimizing lung cancer treatment. *Nature reviews. Clinical oncology*. 9:144-155.
- Puissant, A., S.M. Frumm, G. Alexe, C.F. Bassil, J. Qi, Y.H. Chanthery, E.A. Nekritz, R. Zeid, W.C. Gustafson, P. Greninger, M.J. Garnett, U. McDermott, C.H. Benes, A.L. Kung, W.A. Weiss, J.E. Bradner, and K. Stegmaier. 2013. Targeting MYCN in neuroblastoma by BET bromodomain inhibition. *Cancer discovery*. 3:308-323.
- Pulling, L.C., K.K. Divine, D.M. Klinge, F.D. Gilliland, T. Kang, A.G. Schwartz, T.J. Bocklage, and S.A. Belinsky. 2003. Promoter hypermethylation of the O6-methylguanine-DNA methyltransferase gene: more common in lung adenocarcinomas from never-smokers than smokers and associated with tumor progression. *Cancer research*. 63:4842-4848.
- Rabbitts, T.H. 1994. Chromosomal translocations in human cancer. *Nature*. 372:143-149.
- Rachez, C., and L.P. Freedman. 2000. Mechanisms of gene regulation by vitamin D(3) receptor: a network of coactivator interactions. *Gene*. 246:9-21.
- Reed, J.C., T. Miyashita, S. Takayama, H.G. Wang, T. Sato, S. Krajewski, C. Aime-Sempe, S. Bodrug, S. Kitada, and M. Hanada. 1996. BCL-2 family proteins: regulators of cell death involved in the pathogenesis of cancer and resistance to therapy. *Journal of cellular biochemistry*. 60:23-32.
- Reik, W., and J. Walter. 2001. Genomic imprinting: parental influence on the genome. *Nature reviews. Genetics*. 2:21-32.
- Reinke, H., and W. Horz. 2003. Histones are first hyperacetylated and then lose contact with the activated PHO5 promoter. *Molecular cell*. 11:1599-1607.
- Reynold, N., S. Rousseaux, and S. Khochbin. 2010a. [Epigenetic perturbations and cancer: innovative therapeutic strategies against cancer]. *Bulletin du cancer*. 97:1265-1274.
- Reynold, N., B.E. Schwartz, M. Delvecchio, K. Sadoul, D. Meyers, C. Mukherjee, C. Caron, H. Kimura, S. Rousseaux, P.A. Cole, D. Panne, C.A. French, and S. Khochbin. 2010b. Oncogenesis by sequestration of CBP/p300 in transcriptionally inactive hyperacetylated chromatin domains. *The EMBO journal*. 29:2943-2952.
- Reynolds, C., K.T. Montone, C.M. Powell, J.E. Tomaszewski, and C.V. Clevenger. 1997. Expression of prolactin and its receptor in human breast carcinoma. *Endocrinology*. 138:5555-5560.
- Rhee, I., K.E. Bachman, B.H. Park, K.W. Jair, R.W. Yen, K.E. Schuebel, H. Cui, A.P. Feinberg, C. Lengauer, K.W. Kinzler, S.B. Baylin, and B. Vogelstein. 2002. DNMT1 and DNMT3b cooperate to silence genes in human cancer cells. *Nature*. 416:552-556.
- Richmond, T.J., and C.A. Davey. 2003. The structure of DNA in the nucleosome core. *Nature*. 423:145-150.
- Riddle, O.B., RW ; Dykshorn, SW 1933. The preparation, identification and assay of prolactin – a hormone of the anterior pituitary. *American Journal of Physiology*. 105:191-216.
- Riely, G.J., M.G. Kris, D. Rosenbaum, J. Marks, A. Li, D.A. Chitale, K. Nafa, E.R. Riedel, M. Hsu, W. Pao, V.A. Miller, and M. Ladanyi. 2008. Frequency and distinctive spectrum of KRAS mutations in never smokers with lung adenocarcinoma. *Clinical cancer research : an official journal of the American Association for Cancer Research*. 14:5731-5734.
- Risinger, J.I., G.V. Chandramouli, G.L. Maxwell, M. Custer, S. Pack, D. Loukinov, O. Aprelikova, T. Litzi, D.S. Schrupp, S.K. Murphy, A. Berchuck, V. Lobanenko, and J.C. Barrett. 2007. Global expression analysis of cancer/testis genes in uterine cancers reveals a high incidence of BORIS expression. *Clinical cancer research : an official journal of the American Association for Cancer Research*. 13:1713-1719.
- Ritzenthaler, J.D., S. Han, and J. Roman. 2008. Stimulation of lung carcinoma cell growth by fibronectin-integrin signalling. *Molecular bioSystems*. 4:1160-1169.
- Rivera, M.P., F. Detterbeck, A.C. Mehta, and P. American College of Chest. 2003. Diagnosis of lung cancer: the guidelines. *Chest*. 123:129S-136S.
- Rodenhuis, S., and R.J. Slebos. 1992. Clinical significance of ras oncogene activation in human lung cancer. *Cancer research*. 52:2665S-2669S.
- Rodin, S.N., and A.S. Rodin. 2000. Human lung cancer and p53: the interplay between mutagenesis and selection. *Proceedings of the National Academy of Sciences of the United States of America*. 97:12244-12249.
- Rodin, S.N., and A.S. Rodin. 2005. Origins and selection of p53 mutations in lung carcinogenesis. *Seminars in cancer biology*. 15:103-112.

- Rose-Hellekant, T.A., L.M. Arendt, M.D. Schroeder, K. Gilchrist, E.P. Sandgren, and L.A. Schuler. 2003. Prolactin induces ERalpha-positive and ERalpha-negative mammary cancer in transgenic mice. *Oncogene*. 22:4664-4674.
- Rosen, J.M., D.L. O'Neal, J.E. McHugh, and J.P. Comstock. 1978. Progesterone-mediated inhibition of casein mRNA and polysomal casein synthesis in the rat mammary gland during pregnancy. *Biochemistry*. 17:290-297.
- Roskoski, R., Jr. 2013. Anaplastic lymphoma kinase (ALK): structure, oncogenic activation, and pharmacological inhibition. *Pharmacological research : the official journal of the Italian Pharmacological Society*. 68:68-94.
- Ross, M.T., D.V. Grafham, A.J. Coffey, S. Scherer, K. McLay, D. Muzny, M. Platzer, G.R. Howell, C. Burrows, C.P. Bird, A. Frankish, F.L. Lovell, K.L. Howe, J.L. Ashurst, R.S. Fulton, R. Sudbrak, G. Wen, M.C. Jones, M.E. Hurles, T.D. Andrews, C.E. Scott, S. Searle, J. Ramser, A. Whittaker, R. Deadman, N.P. Carter, S.E. Hunt, R. Chen, A. Cree, P. Gunaratne, P. Havlak, A. Hodgson, M.L. Metzker, S. Richards, G. Scott, D. Steffen, E. Sodergren, D.A. Wheeler, K.C. Worley, R. Ainscough, K.D. Ambrose, M.A. Ansari-Lari, S. Aradhya, R.I. Ashwell, A.K. Babbage, C.L. Bagguley, A. Ballabio, R. Banerjee, G.E. Barker, K.F. Barlow, I.P. Barrett, K.N. Bates, D.M. Beare, H. Beasley, O. Beasley, A. Beck, G. Bethel, K. Blechschmidt, N. Brady, S. Bray-Allen, A.M. Bridgeman, A.J. Brown, M.J. Brown, D. Bonnin, E.A. Bruford, C. Buhay, P. Burch, D. Burford, J. Burgess, W. Burrill, J. Burton, J.M. Bye, C. Carder, L. Carrel, J. Chako, J.C. Chapman, D. Chavez, E. Chen, G. Chen, Y. Chen, Z. Chen, C. Chinault, A. Ciccodicola, S.Y. Clark, G. Clarke, C.M. Clee, S. Clegg, K. Clerc-Blankenburg, K. Clifford, V. Cobley, C.G. Cole, J.S. Conquer, N. Corby, R.E. Connor, R. David, J. Davies, C. Davis, J. Davis, O. Delgado, D. Deshazo, et al. 2005. The DNA sequence of the human X chromosome. *Nature*. 434:325-337.
- Rosti, G., G. Bevilacqua, P. Bidoli, L. Portalone, A. Santo, and G. Genestreti. 2006. Small cell lung cancer. *Annals of oncology : official journal of the European Society for Medical Oncology / ESMO*. 17 Suppl 2:ii5-10.
- Rouet, V., R.L. Bogorad, C. Kayser, K. Kessal, C. Genestie, A. Bardier, D.R. Grattan, B. Kelder, J.J. Kopchick, P.A. Kelly, and V. Goffin. 2010. Local prolactin is a target to prevent expansion of basal/stem cells in prostate tumors. *Proceedings of the National Academy of Sciences of the United States of America*. 107:15199-15204.
- Rousseaux, S., F. Boussouar, J. Gaucher, N. Reynoird, E. Montellier, S. Curtet, A.L. Vitte, and S. Khochbin. 2011. Molecular models for post-meiotic male genome reprogramming. *Systems biology in reproductive medicine*. 57:50-53.
- Rousseaux, S., and S. Khochbin. 2009. New hypotheses for large-scale epigenome alterations in somatic cancer cells: a role for male germ-cell-specific regulators. *Epigenomics*. 1:153-161.
- Rousseaux, S., N. Reynoird, E. Escoffier, J. Thevenon, C. Caron, and S. Khochbin. 2008. Epigenetic reprogramming of the male genome during gametogenesis and in the zygote. *Reproductive biomedicine online*. 16:492-503.
- Rousseaux, S.D., A.; Jacquiau, B.; Vitte, A. L.; Vesin, A.; Nagy-Mignotte, H.; Moro-Sibilot, D.; Brichon, P. Y.; Lantuejoul, S.; Hainaut, P.; Laffaire, J.; de Reyniès, A.; Beer, D. G.; Timsit, J. F.; Brambilla, C.; Brambilla, E.; Khochbin, S. . 2013. Ectopic activation of germline and placental genes identifies aggressive metastasis-prone lung cancers. *Science Translational Medicine*. In press.
- Roz, L., M. Gramegna, H. Ishii, C.M. Croce, and G. Sozzi. 2002. Restoration of fragile histidine triad (FHIT) expression induces apoptosis and suppresses tumorigenicity in lung and cervical cancer cell lines. *Proceedings of the National Academy of Sciences of the United States of America*. 99:3615-3620.
- Ruden, M., and N. Puri. 2013. Novel anticancer therapeutics targeting telomerase. *Cancer treatment reviews*. 39:444-456.
- Rudolph, M.C., J.L. McManaman, L. Hunter, T. Phang, and M.C. Neville. 2003. Functional development of the mammary gland: use of expression profiling and trajectory clustering to reveal changes in gene expression during pregnancy, lactation, and involution. *Journal of mammary gland biology and neoplasia*. 8:287-307.
- Rudolph, M.C., J.L. McManaman, T. Phang, T. Russell, D.J. Kominsky, N.J. Serkova, T. Stein, S.M. Anderson, and M.C. Neville. 2007. Metabolic regulation in the lactating mammary gland: a lipid synthesizing machine. *Physiological genomics*. 28:323-336.
- Russell, L.D.E., R. A. ; Sinha Hikim, A. P. ; Clegg, E. D. 1990. Histological and Histopathological Evaluation of the Testis. . Cache River Press, Clearwater, Florida.
- Ryan, R.F., D.C. Schultz, K. Ayyanathan, P.B. Singh, J.R. Friedman, W.J. Fredericks, and F.J. Rauscher, 3rd. 1999. KAP-1 corepressor protein interacts and colocalizes with heterochromatic and euchromatic HP1 proteins: a potential role for Kruppel-associated box-zinc finger proteins in heterochromatin-mediated gene silencing. *Molecular and cellular biology*. 19:4366-4378.
- Ryberg, D., E. Kure, S. Lystad, V. Skaug, L. Stangeland, I. Mercy, A.L. Borresen, and A. Haugen. 1994. p53 mutations in lung tumors: relationship to putative susceptibility markers for cancer. *Cancer research*. 54:1551-1555.
- Sadoul, K., J. Wang, B. Diagouraga, and S. Khochbin. 2011. The tale of protein lysine acetylation in the cytoplasm. *Journal of biomedicine & biotechnology*. 2011:970382.
- Sagawa, Y., H. Nishi, K. Isaka, A. Fujito, and M. Takayama. 2001. The correlation of TERT expression with c-myc expression in cervical cancer. *Cancer letters*. 168:45-50.
- Saha, A., J. Wittmeyer, and B.R. Cairns. 2005. Chromatin remodeling through directional DNA translocation from an internal nucleosomal site. *Nature structural & molecular biology*. 12:747-755.
- Sahin, U., O. Tureci, Y.T. Chen, G. Seitz, C. Villena-Heinsen, L.J. Old, and M. Pfreundschuh. 1998. Expression of multiple cancer/testis (CT) antigens in breast cancer and melanoma: basis for polyvalent CT vaccine strategies. *International journal of cancer. Journal international du cancer*. 78:387-389.
- Sahin, U., O. Tureci, H. Schmitt, B. Cochlovius, T. Johannes, R. Schmits, F. Stenner, G. Luo, I. Schobert, and M. Pfreundschuh. 1995. Human neoplasms elicit multiple specific immune responses in the autologous host. *Proceedings of the National Academy of Sciences of the United States of America*. 92:11810-11813.

- Salcido, C.D., A. Larochelle, B.J. Taylor, C.E. Dunbar, and L. Varticovski. 2010. Molecular characterisation of side population cells with cancer stem cell-like characteristics in small-cell lung cancer. *British journal of cancer*. 102:1636-1644.
- Sanders, H.R., and M. Albitar. 2010. Somatic mutations of signaling genes in non-small-cell lung cancer. *Cancer genetics and cytogenetics*. 203:7-15.
- Sard, L., P. Accornero, S. Tornielli, D. Delia, G. Bunone, M. Campiglio, M.P. Colombo, M. Gramegna, C.M. Croce, M.A. Pierotti, and G. Sozzi. 1999. The tumor-suppressor gene FHIT is involved in the regulation of apoptosis and in cell cycle control. *Proceedings of the National Academy of Sciences of the United States of America*. 96:8489-8492.
- Sarraf, S.A., and I. Stancheva. 2004. Methyl-CpG binding protein MBD1 couples histone H3 methylation at lysine 9 by SETDB1 to DNA replication and chromatin assembly. *Molecular cell*. 15:595-605.
- Sasaki, H., S. Moriyama, Y. Nakashima, Y. Kobayashi, M. Kiriyama, I. Fukai, Y. Yamakawa, and Y. Fujii. 2004. Histone deacetylase 1 mRNA expression in lung cancer. *Lung cancer*. 46:171-178.
- Sasaki, T., S.J. Rodig, L.R. Chirieac, and P.A. Janne. 2010. The biology and treatment of EML4-ALK non-small cell lung cancer. *European journal of cancer*. 46:1773-1780.
- Sato, M., D.S. Shames, A.F. Gazdar, and J.D. Minna. 2007. A translational view of the molecular pathogenesis of lung cancer. *Journal of thoracic oncology : official publication of the International Association for the Study of Lung Cancer*. 2:327-343.
- Sawan, C., T. Vaissiere, R. Murr, and Z. Herceg. 2008. Epigenetic drivers and genetic passengers on the road to cancer. *Mutation research*. 642:1-13.
- Scanlan, M.J., C.M. Gordon, B. Williamson, S.Y. Lee, Y.T. Chen, E. Stockert, A. Jungbluth, G. Ritter, D. Jager, E. Jager, A. Knuth, and L.J. Old. 2002a. Identification of cancer/testis genes by database mining and mRNA expression analysis. *International journal of cancer. Journal international du cancer*. 98:485-492.
- Scanlan, M.J., A.O. Gure, A.A. Jungbluth, L.J. Old, and Y.T. Chen. 2002b. Cancer/testis antigens: an expanding family of targets for cancer immunotherapy. *Immunological reviews*. 188:22-32.
- Schally, A.V., A. Kuroshima, Y. Ishida, T.W. Redding, and C.Y. Bowers. 1965. The Presence of Prolactin Inhibiting Factor (Pif) in Extracts of Beef, Sheep and Pig Hypothalami. *Proceedings of the Society for Experimental Biology and Medicine. Society for Experimental Biology and Medicine*. 118:350-352.
- Schmitt-Ney, M., W. Doppler, R.K. Ball, and B. Groner. 1991. Beta-casein gene promoter activity is regulated by the hormone-mediated relief of transcriptional repression and a mammary-gland-specific nuclear factor. *Molecular and cellular biology*. 11:3745-3755.
- Schones, D.E., and K. Zhao. 2008. Genome-wide approaches to studying chromatin modifications. *Nature reviews. Genetics*. 9:179-191.
- Schroeder, M.D., J. Symowicz, and L.A. Schuler. 2002. PRL modulates cell cycle regulators in mammary tumor epithelial cells. *Molecular endocrinology*. 16:45-57.
- Schultz, D.C., J.R. Friedman, and F.J. Rauscher, 3rd. 2001. Targeting histone deacetylase complexes via KRAB-zinc finger proteins: the PHD and bromodomains of KAP-1 form a cooperative unit that recruits a novel isoform of the Mi-2alpha subunit of NuRD. *Genes & development*. 15:428-443.
- Schwartz, B.E., M.D. Hofer, M.E. Lemieux, D.E. Bauer, M.J. Cameron, N.H. West, E.S. Agoston, N. Reynoird, S. Khochbin, T.A. Ince, A. Christie, K.A. Janeway, S.O. Vargas, A.R. Perez-Atayde, J.C. Aster, S.E. Sallan, A.L. Kung, J.E. Bradner, and C.A. French. 2011. Differentiation of NUT midline carcinoma by epigenomic reprogramming. *Cancer research*. 71:2686-2696.
- Selamat, S.A., J.S. Galler, A.D. Joshi, M.N. Fyfe, M. Campan, K.D. Siegmund, K.M. Kerr, and I.A. Laird-Offringa. 2011. DNA methylation changes in atypical adenomatous hyperplasia, adenocarcinoma in situ, and lung adenocarcinoma. *PloS one*. 6:e21443.
- Seligson, D.B., S. Horvath, M.A. McBrien, V. Mah, H. Yu, S. Tze, Q. Wang, D. Chia, L. Goodglick, and S.K. Kurdastani. 2009. Global levels of histone modifications predict prognosis in different cancers. *The American journal of pathology*. 174:1619-1628.
- Semprini, S., S. Friedrichsen, C.V. Harper, J.R. McNeilly, A.D. Adamson, D.G. Spiller, N. Kotelevtseva, G. Brooker, D.G. Brownstein, A.S. McNeilly, M.R. White, J.R. Davis, and J.J. Mullins. 2009. Real-time visualization of human prolactin alternate promoter usage in vivo using a double-transgenic rat model. *Molecular endocrinology*. 23:529-538.
- Sequist, L.V., D.W. Bell, T.J. Lynch, and D.A. Haber. 2007. Molecular predictors of response to epidermal growth factor receptor antagonists in non-small-cell lung cancer. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology*. 25:587-595.
- Sethi, B.K., G.V. Chanukya, and V.S. Nagesh. 2012. Prolactin and cancer: Has the orphan finally found a home? *Indian journal of endocrinology and metabolism*. 16:S195-198.
- Shang, E., H.D. Nickerson, D. Wen, X. Wang, and D.J. Wolgemuth. 2007. The first bromodomain of Brdt, a testis-specific member of the BET sub-family of double-bromodomain-containing proteins, is essential for male germ cell differentiation. *Development*. 134:3507-3515.
- Shaw-Bruha, C.M., S.J. Pirruccello, and J.D. Shull. 1997. Expression of the prolactin gene in normal and neoplastic human breast tissues and human mammary cell lines: promoter usage and alternative mRNA splicing. *Breast cancer research and treatment*. 44:243-253.

- Shaw, A.T., and B. Solomon. 2011. Targeting anaplastic lymphoma kinase in lung cancer. *Clinical cancer research : an official journal of the American Association for Cancer Research*. 17:2081-2086.
- Shi, Y., F. Lan, C. Matson, P. Mulligan, J.R. Whetstone, P.A. Cole, R.A. Casero, and Y. Shi. 2004. Histone demethylation mediated by the nuclear amine oxidase homolog LSD1. *Cell*. 119:941-953.
- Shigematsu, Y., T. Hanagiri, H. Shiota, K. Kuroda, T. Baba, M. Mizukami, T. So, Y. Ichiki, M. Yasuda, T. So, M. Takenoyama, and K. Yasumoto. 2010. Clinical significance of cancer/testis antigens expression in patients with non-small cell lung cancer. *Lung cancer*. 68:105-110.
- Siegfried, J.M., A.L. Davis, J.R. Testa, J.D. Hunt, R.J. Keenan, S.A. Yousem, P.S. Ritter, P.F. Ferson, and R.J. Landreneau. 1994. Ability to culture resectable non-small cell lung carcinomas is correlated with recurrence. *The Annals of thoracic surgery*. 58:662-666; discussion 667.
- Simpson, A.J., O.L. Caballero, A. Jungbluth, Y.T. Chen, and L.J. Old. 2005. Cancer/testis antigens, gametogenesis and cancer. *Nature reviews. Cancer*. 5:615-625.
- Singh, B.N., G. Zhang, Y.L. Hwa, J. Li, S.C. Dowdy, and S.W. Jiang. 2010. Nonhistone protein acetylation as cancer therapy targets. *Expert review of anticancer therapy*. 10:935-954.
- Siprashvili, Z., G. Sozzi, L.D. Barnes, P. McCue, A.K. Robinson, V. Eryomin, L. Sard, E. Tagliabue, A. Greco, L. Fusetti, G. Schwartz, M.A. Pierotti, C.M. Croce, and K. Huebner. 1997. Replacement of Fhit in cancer cells suppresses tumorigenicity. *Proceedings of the National Academy of Sciences of the United States of America*. 94:13771-13776.
- Smale, S.T., and J.T. Kadonaga. 2003. The RNA polymerase II core promoter. *Annual review of biochemistry*. 72:449-479.
- Soda, M., Y.L. Choi, M. Enomoto, S. Takada, Y. Yamashita, S. Ishikawa, S. Fujiwara, H. Watanabe, K. Kurashina, H. Hatanaka, M. Bando, S. Ohno, Y. Ishikawa, H. Aburatani, T. Niki, Y. Sohara, Y. Sugiyama, and H. Mano. 2007. Identification of the transforming EML4-ALK fusion gene in non-small-cell lung cancer. *Nature*. 448:561-566.
- Soda, M., S. Takada, K. Takeuchi, Y.L. Choi, M. Enomoto, T. Ueno, H. Haruta, T. Hamada, Y. Yamashita, Y. Ishikawa, Y. Sugiyama, and H. Mano. 2008. A mouse model for EML4-ALK-positive lung cancer. *Proceedings of the National Academy of Sciences of the United States of America*. 105:19893-19897.
- Sorenson, G.D., O.S. Pettengill, T. Brinck-Johnsen, C.C. Cate, and L.H. Maurer. 1981. Hormone production by cultures of small-cell carcinoma of the lung. *Cancer*. 47:1289-1296.
- Sozzi, G., U. Pastorino, L. Moiraghi, E. Tagliabue, F. Pezzella, C. Ghirelli, S. Tornielli, L. Sard, K. Huebner, M.A. Pierotti, C.M. Croce, and S. Pilotti. 1998. Loss of FHIT function in lung cancer and preinvasive bronchial lesions. *Cancer research*. 58:5032-5037.
- Sozzi, G., S. Tornielli, E. Tagliabue, L. Sard, F. Pezzella, U. Pastorino, F. Minoletti, S. Pilotti, C. Ratcliffe, M.L. Veronese, P. Goldstraw, K. Huebner, C.M. Croce, and M.A. Pierotti. 1997. Absence of Fhit protein in primary lung tumors and cell lines with FHIT gene abnormalities. *Cancer research*. 57:5207-5212.
- Sozzi, G., M.L. Veronese, M. Negrini, R. Baffa, M.G. Cotticelli, H. Inoue, S. Tornielli, S. Pilotti, L. De Gregorio, U. Pastorino, M.A. Pierotti, M. Ohta, K. Huebner, and C.M. Croce. 1996. The FHIT gene 3p14.2 is abnormal in lung cancer. *Cell*. 85:17-26.
- Spiro, S.G., and G.A. Silvestri. 2005. One hundred years of lung cancer. *American journal of respiratory and critical care medicine*. 172:523-529.
- Sripathy, S.P., J. Stevens, and D.C. Schultz. 2006. The KAP1 corepressor functions to coordinate the assembly of de novo HP1-demarcated microenvironments of heterochromatin required for KRAB zinc finger protein-mediated transcriptional repression. *Molecular and cellular biology*. 26:8623-8638.
- Stattin, P., S. Rinaldi, U.H. Stenman, E. Riboli, G. Hallmans, A. Bergh, and R. Kaaks. 2001. Plasma prolactin and prostate cancer risk: A prospective study. *International journal of cancer. Journal international du cancer*. 92:463-465.
- Sterner, D.E., and S.L. Berger. 2000. Acetylation of histones and transcription-related factors. *Microbiology and molecular biology reviews : MMBR*. 64:435-459.
- Sternlicht, M.D. 2006. Key stages in mammary gland development: the cues that regulate ductal branching morphogenesis. *Breast cancer research : BCR*. 8:201.
- Strahl, B.D., and C.D. Allis. 2000. The language of covalent histone modifications. *Nature*. 403:41-45.
- Strasser, A., D.C. Huang, and D.L. Vaux. 1997. The role of the bcl-2/ced-9 gene family in cancer and general implications of defects in cell death control for tumorigenesis and resistance to chemotherapy. *Biochimica et biophysica acta*. 1333:F151-178.
- Stratton, M.R. 2011. Exploring the genomes of cancer cells: progress and promise. *Science*. 331:1553-1558.
- Stricker, P.G., R 1928. Action du lobe antérieur de l'hypophyse sur la montée laiteuse. *Comptes Rendus des Séances de la Société de Biologie et de ses Filiales*
- Sturla, S.J. 2007. DNA adduct profiles: chemical approaches to addressing the biological impact of DNA damage from small molecules. *Current opinion in chemical biology*. 11:293-299.
- Subramanian, J., and R. Govindan. 2007. Lung cancer in never smokers: a review. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology*. 25:561-570.
- Sun, S., J.H. Schiller, and A.F. Gazdar. 2007a. Lung cancer in never smokers--a different disease. *Nature reviews. Cancer*. 7:778-790.
- Sun, S., J.H. Schiller, M. Spinola, and J.D. Minna. 2007b. New molecularly targeted therapies for lung cancer. *The Journal of clinical investigation*. 117:2740-2750.
- Suryaraja, R., M. Anitha, K. Anbarasu, G. Kumari, and S. Mahalingam. 2013. The E3 ubiquitin ligase Itch regulates tumor suppressor protein RASSF5/NORE1 stability in an acetylation-dependent manner. *Cell death & disease*. 4:e565.

- Suzuki, M., and I. Yoshino. 2010. Aberrant methylation in non-small cell lung cancer. *Surgery today*. 40:602-607.
- Szostak, J.W. 2010. DNA ends: just the beginning (Nobel lecture). *Angewandte Chemie*. 49:7386-7404.
- Szyf, M., P. Pakneshan, and S.A. Rabbani. 2004. DNA demethylation and cancer: therapeutic implications. *Cancer letters*. 211:133-143.
- Tabruyn, S.P., N.Q. Nguyen, A.M. Cornet, J.A. Martial, and I. Struman. 2005. The antiangiogenic factor, 16-kDa human prolactin, induces endothelial cell cycle arrest by acting at both the Go-G1 and the G2-M phases. *Molecular endocrinology*. 19:1932-1942.
- Tabruyn, S.P., C.M. Sorlet, F. Rentier-Delrue, V. Bours, R.I. Weiner, J.A. Martial, and I. Struman. 2003. The antiangiogenic factor 16K human prolactin induces caspase-dependent apoptosis by a mechanism that requires activation of nuclear factor-kappaB. *Molecular endocrinology*. 17:1815-1823.
- Taggart, R.T., T.K. Mohandas, and G.I. Bell. 1987. Assignment of the human preprogastricsin (PGC) to chromosome 6 and regional localization of PGC (6pter-p21.1), prolactin PRL (6pter-p21.1). *Cytogenet. Cell Genet*. 46:701-702.
- Tajima, K., Y. Obata, H. Tamaki, M. Yoshida, Y.T. Chen, M.J. Scanlan, L.J. Old, H. Kuwano, T. Takahashi, T. Takahashi, and T. Mitsudomi. 2003. Expression of cancer/testis (CT) antigens in lung cancer. *Lung cancer*. 42:23-33.
- Takahashi, T., M.M. Nau, I. Chiba, M.J. Birrer, R.K. Rosenberg, M. Vinocour, M. Levitt, H. Pass, A.F. Gazdar, and J.D. Minna. 1989. p53: a frequent target for genetic abnormalities in lung cancer. *Science*. 246:491-494.
- Talikka, M., N. Sierro, N.V. Ivanov, N. Chaudhary, M.J. Peck, J. Hoeng, C.R. Coggins, and M.C. Peitsch. 2012. Genomic impact of cigarette smoke, with application to three smoking-related diseases. *Critical reviews in toxicology*. 42:877-889.
- Tan, D., G. Deeb, J. Wang, H.K. Slocum, J. Winston, S. Wiseman, A. Beck, S. Sait, T. Anderson, C. Nwogu, N. Ramnath, and G. Loewen. 2003. HER-2/neu protein expression and gene alteration in stage I-IIIa non-small-cell lung cancer: a study of 140 cases using a combination of high throughput tissue microarray, immunohistochemistry, and fluorescent in situ hybridization. *Diagnostic molecular pathology : the American journal of surgical pathology, part B*. 12:201-211.
- Tan, M., H. Luo, S. Lee, F. Jin, J.S. Yang, E. Montellier, T. Buchou, Z. Cheng, S. Rousseaux, N. Rajagopal, Z. Lu, Z. Ye, Q. Zhu, J. Wysocka, Y. Ye, S. Khochbin, B. Ren, and Y. Zhao. 2011. Identification of 67 histone marks and histone lysine crotonylation as a new type of histone modification. *Cell*. 146:1016-1028.
- Taverna, S.D., B.M. Ueberheide, Y. Liu, A.J. Tackett, R.L. Diaz, J. Shabanowitz, B.T. Chait, D.F. Hunt, and C.D. Allis. 2007. Long-distance combinatorial linkage between methylation and acetylation on histone H3 N termini. *Proceedings of the National Academy of Sciences of the United States of America*. 104:2086-2091.
- Testa, J.R., J.M. Siegfried, Z. Liu, J.D. Hunt, M.M. Feder, S. Litwin, J.Y. Zhou, T. Taguchi, and S.M. Keller. 1994. Cytogenetic analysis of 63 non-small cell lung carcinomas: recurrent chromosome alterations amid frequent and widespread genomic upheaval. *Genes, chromosomes & cancer*. 11:178-194.
- Thiagalingam, S., K.H. Cheng, H.J. Lee, N. Mineva, A. Thiagalingam, and J.F. Ponte. 2003. Histone deacetylases: unique players in shaping the epigenetic histone code. *Annals of the New York Academy of Sciences*. 983:84-100.
- Thilly, W.G. 1983. Analysis of chemically induced mutation in single cell populations. *Basic life sciences*. 23:337-378.
- Thomas, M.C., and C.M. Chiang. 2006. The general transcription machinery and general cofactors. *Critical reviews in biochemistry and molecular biology*. 41:105-178.
- Thun, M.J., L.M. Hannan, L.L. Adams-Campbell, P. Boffetta, J.E. Buring, D. Feskanich, W.D. Flanders, S.H. Jee, K. Katanoda, L.N. Kolonel, I.M. Lee, T. Marugame, J.R. Palmer, E. Riboli, T. Sobue, E. Avila-Tang, L.R. Wilkens, and J.M. Samet. 2008. Lung cancer occurrence in never-smokers: an analysis of 13 cohorts and 22 cancer registry studies. *PLoS medicine*. 5:e185.
- Tjian, R. 1996. The biochemistry of transcription in eukaryotes: a paradigm for multisubunit regulatory complexes. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*. 351:491-499.
- Tomio, A., D.J. Schust, K. Kawana, T. Yasugi, Y. Kawana, S. Mahalingaiah, T. Fujii, and Y. Taketani. 2008. Prolactin can modulate CD4+ T-cell response through receptor-mediated alterations in the expression of T-bet. *Immunology and cell biology*. 86:616-621.
- Tootle, T.L., and I. Rebay. 2005. Post-translational modifications influence transcription factor activity: a view from the ETS superfamily. *BioEssays : news and reviews in molecular, cellular and developmental biology*. 27:285-298.
- Tornaletti, S., and G.P. Pfeifer. 1995. Complete and tissue-independent methylation of CpG sites in the p53 gene: implications for mutations in human cancers. *Oncogene*. 10:1493-1499.
- Touraine, P., J.F. Martini, B. Zafrani, J.C. Durand, F. Labaille, C. Malet, A. Nicolas, C. Trivin, M.C. Postel-Vinay, F. Kuttann, and P.A. Kelly. 1998. Increased expression of prolactin receptor gene assessed by quantitative polymerase chain reaction in human breast tumors versus normal breast tissues. *The Journal of clinical endocrinology and metabolism*. 83:667-674.
- Tourani, J.M. Stratégies thérapeutiques devant les cancers bronchiques à petites cellules.
- Tran-Thanh, D., N.C. Arneson, M. Pintilie, A. Deliallisi, K.S. Warren, A. Bane, and S.J. Done. 2011. Amplification of the prolactin receptor gene in mammary lobular neoplasia. *Breast cancer research and treatment*. 128:31-40.
- travail, A.A.f.d.s.d.l.e.e.d. 2009. Pollution par les particules dans l'air ambiant

Synthèse des éléments sanitaires en vue d'un appui à l'élaboration de seuils d'information et d'alerte du public pour les particules dans l'air ambiant. *Rapport d'expertise collective*. Mars 2009.

- Travis, W.D., E. Brambilla, H.K. Muller-Hermelink, and C.C. Harris. 2004. World Health Organization Classification of Tumours. Pathology and Genetics of Tumours of the Lung, Pleura, Thymus and Heart. IARC Press: Lyon.
- Triebel, J., Y. Macotela, G.M. de la Escalera, and C. Clapp. 2011. Prolactin and vasoinhibins: Endogenous players in diabetic retinopathy. *IUBMB life*. 63:806-810.
- Truong, A.T., C. Duez, A. Belayew, A. Renard, R. Pictet, G.I. Bell, and J.A. Martial. 1984. Isolation and characterization of the human prolactin gene. *The EMBO journal*. 3:429-437.
- Tworoger, S.S., and S.E. Hankinson. 2008. Prolactin and breast cancer etiology: an epidemiologic perspective. *Journal of mammary gland biology and neoplasia*. 13:41-53.
- Tworoger, S.S., S.A. Missmer, R.L. Barbieri, W.C. Willett, G.A. Colditz, and S.E. Hankinson. 2005. Plasma sex hormone concentrations and subsequent risk of breast cancer among women using postmenopausal hormones. *Journal of the National Cancer Institute*. 97:595-602.
- Vallin, J., and F. Meslé. 1988. Les causes de décès en France de 1925 à 1978.
- Van Den Broeck, A., E. Brambilla, D. Moro-Sibilot, S. Lantuejoul, C. Brambilla, B. Eymin, S. Khochbin, and S. Gazzeri. 2008. Loss of histone H4K20 trimethylation occurs in preneoplasia and influences prognosis of non-small cell lung cancer. *Clinical cancer research : an official journal of the American Association for Cancer Research*. 14:7237-7245.
- Van Den Broeck, A., P. Ozenne, B. Eymin, and S. Gazzeri. 2010. Lung cancer: a modified epigenome. *Cell adhesion & migration*. 4:107-113.
- van der Bruggen, P., C. Traversari, P. Chomez, C. Lurquin, E. De Plaen, B. Van den Eynde, A. Knuth, and T. Boon. 1991. A gene encoding an antigen recognized by cytolytic T lymphocytes on a human melanoma. *Science*. 254:1643-1647.
- van der Ploeg, L.H., and R.A. Flavell. 1980. DNA methylation in the human gamma delta beta-globin locus in erythroid and nonerythroid tissues. *Cell*. 19:947-958.
- van Gent, D.C., J.H. Hoeijmakers, and R. Kanaar. 2001. Chromosomal stability and the DNA double-stranded break connection. *Nature reviews. Genetics*. 2:196-206.
- Vaporciyan, A.A., K.W. Merriman, F. Ece, J.A. Roth, W.R. Smythe, S.G. Swisher, G.L. Walsh, J.C. Nesbitt, and J.B. Putnam, Jr. 2002. Incidence of major pulmonary morbidity after pneumonectomy: association with timing of smoking cessation. *The Annals of thoracic surgery*. 73:420-425; discussion 425-426.
- Varella-Garcia, M., L. Chen, R.L. Powell, F.R. Hirsch, T.C. Kennedy, R. Keith, Y.E. Miller, J.D. Mitchell, and W.A. Franklin. 2007. Spectral karyotyping detects chromosome damage in bronchial cells of smokers and patients with cancer. *American journal of respiratory and critical care medicine*. 176:505-512.
- Vatolin, S., Z. Abdullaev, S.D. Pack, P.T. Flanagan, M. Custer, D.I. Loukinov, E. Pugacheva, J.A. Hong, H. Morse, 3rd, D.S. Schrupp, J.I. Risinger, J.C. Barrett, and V.V. Lobanenko. 2005. Conditional expression of the CTCF-paralogous transcriptional factor BORIS in normal cells results in demethylation and derepression of MAGE-A1 and reactivation of other cancer-testis genes. *Cancer research*. 65:7751-7762.
- Verschuere, K., J.E. Remacle, C. Collart, H. Kraft, B.S. Baker, P. Tylzanowski, L. Nelles, G. Wuytens, M.T. Su, R. Bodmer, J.C. Smith, and D. Huylebroeck. 1999. SIP1, a novel zinc finger/homeodomain repressor, interacts with Smad proteins and binds to 5'-CACCT sequences in candidate target genes. *The Journal of biological chemistry*. 274:20489-20498.
- Wade, P.A. 2001. Methyl CpG-binding proteins and transcriptional repression. *BioEssays : news and reviews in molecular, cellular and developmental biology*. 23:1131-1137.
- Wakao, H., F. Gouilleux, and B. Groner. 1994. Mammary gland factor (MGF) is a novel member of the cytokine regulated transcription factor gene family and confers the prolactin response. *The EMBO journal*. 13:2182-2191.
- Walker, A.M. 1994. Phosphorylated and nonphosphorylated Prolactin isoforms. *Trends in endocrinology and metabolism: TEM*. 5:195-200.
- Wang, D.Y., K.A. Stepniowska, D.S. Allen, I.S. Fentiman, R.D. Bulbrook, H.G. Kwa, B.L. De Stavola, and M.J. Reed. 1995. Serum prolactin levels and their relationship to survival in women with operable breast cancer. *Journal of clinical epidemiology*. 48:959-968.
- Wang, H.B., and Y. Zhang. 2001. Miz, an auto-antigen for dermatomyositis, is an ATP-dependent nucleosome remodeling factor. *Nucleic acids research*. 29:2517-2521.
- Wang, J., A. Emadali, A. Le Bescont, M. Callanan, S. Rousseaux, and S. Khochbin. 2011. Induced malignant genome reprogramming in somatic cells by testis-specific factors. *Biochimica et biophysica acta*. 1809:221-225.
- Wang, J., L.Y. Xie, S. Allan, D. Beach, and G.J. Hannon. 1998. Myc activates telomerase. *Genes & development*. 12:1769-1774.
- Watanabe, H., K. Soejima, H. Yasuda, I. Kawada, I. Nakachi, S. Yoda, K. Naoki, and A. Ishizaka. 2008. Deregulation of histone lysine methyltransferases contributes to oncogenic transformation of human bronchoepithelial cells. *Cancer cell international*. 8:15.
- Watanabe, K., C.A. Kessler, C.J. Bachurski, Y. Kanda, B.D. Richardson, J. Stanek, S. Handwerger, and A.K. Brar. 2001. Identification of a decidua-specific enhancer on the human prolactin gene with two critical activator protein 1 (AP-1) binding sites. *Molecular endocrinology*. 15:638-653.
- Watson, C.J., and T.G. Burdon. 1996. Prolactin signal transduction mechanisms in the mammary gland: the role of the Jak/Stat pathway. *Reviews of reproduction*. 1:1-5.
- Watson, J.D., Baker T.A., Bell S.P., Gann A., Levine M., Losick R. . 2006. Molecular Biology of the Gene. Pearson Higher Ed USA.

- Watt, F., and P.L. Molloy. 1988. Cytosine methylation prevents binding to DNA of a HeLa cell transcription factor required for optimal expression of the adenovirus major late promoter. *Genes & development*. 2:1136-1143.
- Wei, Q., L. Cheng, C.I. Amos, L.E. Wang, Z. Guo, W.K. Hong, and M.R. Spitz. 2000. Repair of tobacco carcinogen-induced DNA adducts and lung cancer risk: a molecular epidemiologic study. *Journal of the National Cancer Institute*. 92:1764-1772.
- Weir, B.A., M.S. Woo, G. Getz, S. Perner, L. Ding, R. Beroukhi, W.M. Lin, M.A. Province, A. Kraja, L.A. Johnson, K. Shah, M. Sato, R.K. Thomas, J.A. Barletta, I.B. Borecki, S. Broderick, A.C. Chang, D.Y. Chiang, L.R. Chirieac, J. Cho, Y. Fujii, A.F. Gazdar, T. Giordano, H. Greulich, M. Hanna, B.E. Johnson, M.G. Kris, A. Lash, L. Lin, N. Lindeman, E.R. Mardis, J.D. McPherson, J.D. Minna, M.B. Morgan, M. Nadel, M.B. Orringer, J.R. Osborne, B. Ozenberger, A.H. Ramos, J. Robinson, J.A. Roth, V. Rusch, H. Sasaki, F. Shepherd, C. Sougnez, M.R. Spitz, M.S. Tsao, D. Twomey, R.G. Verhaak, G.M. Weinstock, D.A. Wheeler, W. Winckler, A. Yoshizawa, S. Yu, M.F. Zakowski, Q. Zhang, D.G. Beer, Wistuba, II, M.A. Watson, L.A. Garraway, M. Ladanyi, W.D. Travis, W. Pao, M.A. Rubin, S.B. Gabriel, R.A. Gibbs, H.E. Varmus, R.K. Wilson, E.S. Lander, and M. Meyerson. 2007. Characterizing the cancer genome in lung adenocarcinoma. *Nature*. 450:893-898.
- Welsch, C.W., and H. Nagasawa. 1977. Prolactin and murine mammary tumorigenesis: a review. *Cancer research*. 37:951-963.
- Welsh, J.S., K. Bradley, K.J. Ruchala, T.R. Mackie, R. Manon, R. Patel, P. Wiederholt, M. Lock, S. Hui, and M.P. Mehta. 2004. Megavoltage computed tomography imaging: a potential tool to guide and improve the delivery of thoracic radiation therapy. *Clinical lung cancer*. 5:303-306.
- Wen, J., J. Fu, W. Zhang, and M. Guo. 2011. Genetic and epigenetic changes in lung carcinoma and their clinical implications. *Modern pathology : an official journal of the United States and Canadian Academy of Pathology, Inc*. 24:932-943.
- Wessely, O., E.M. Deiner, H. Beug, and M. von Lindern. 1997. The glucocorticoid receptor is a key regulator of the decision between self-renewal and differentiation in erythroid progenitors. *The EMBO journal*. 16:267-280.
- West, L., S.J. Vidwans, N.P. Campbell, J. Shrager, G.R. Simon, R. Bueno, P.A. Dennis, G.A. Otterson, and R. Salgia. 2012. A novel classification of lung cancer into molecular subtypes. *PloS one*. 7:e31906.
- Westin, G., and W. Schaffner. 1988. A zinc-responsive factor interacts with a metal-regulated enhancer element (MRE) of the mouse metallothionein-I gene. *The EMBO journal*. 7:3763-3770.
- Weston, A., J.C. Willey, R. Modali, H. Sugimura, E.M. McDowell, J. Resau, B. Light, A. Haugen, D.L. Mann, B.F. Trump, and et al. 1989. Differential DNA sequence deletions from chromosomes 3, 11, 13, and 17 in squamous-cell carcinoma, large-cell carcinoma, and adenocarcinoma of the human lung. *Proceedings of the National Academy of Sciences of the United States of America*. 86:5099-5103.
- Whang-Peng, J., C.S. Kao-Shan, E.C. Lee, P.A. Bunn, D.N. Carney, A.F. Gazdar, and J.D. Minna. 1982. Specific chromosome defect associated with human small-cell lung cancer; deletion 3p(14-23). *Science*. 215:181-182.
- Whetstone, J.R., A. Nottke, F. Lan, M. Huarte, S. Smolnikov, Z. Chen, E. Spooner, E. Li, G. Zhang, M. Colaiacovo, and Y. Shi. 2006. Reversal of histone lysine trimethylation by the JMJD2 family of histone demethylases. *Cell*. 125:467-481.
- WHO, I.A.f.R.o.C. 2002. Tobacco Smoke and Involuntary Smoking - IARC Monographs on the Evaluation of Carcinogenic Risks to Humans. Vol. 83. WHO, International Agency for Research on Cancer.
- Williamson, R. 1970. Properties of rapidly labelled deoxyribonucleic acid fragments isolated from the cytoplasm of primary cultures of embryonic mouse liver cells. *Journal of molecular biology*. 51:157-168.
- Wistuba, II, S. Lam, C. Behrens, A.K. Virmani, K.M. Fong, J. LeRiche, J.M. Samet, S. Srivastava, J.D. Minna, and A.F. Gazdar. 1997. Molecular damage in the bronchial epithelium of current and former smokers. *Journal of the National Cancer Institute*. 89:1366-1373.
- Witschi, H. 2001. A short history of lung cancer. *Toxicological sciences : an official journal of the Society of Toxicology*. 64:4-6.
- Witta, S. 2012. Histone deacetylase inhibitors in non-small-cell lung cancer. *Journal of thoracic oncology : official publication of the International Association for the Study of Lung Cancer*. 7:5404-406.
- Wolffe, A.P., and M.A. Matzke. 1999. Epigenetics: regulation through repression. *Science*. 286:481-486.
- Wu, K.J., C. Grandori, M. Amacker, N. Simon-Vermot, A. Polack, J. Lingner, and R. Dalla-Favera. 1999. Direct activation of TERT transcription by c-MYC. *Nature genetics*. 21:220-224.
- Wu, L.P., X. Wang, L. Li, Y. Zhao, S. Lu, Y. Yu, W. Zhou, X. Liu, J. Yang, Z. Zheng, H. Zhang, J. Feng, Y. Yang, H. Wang, and W.G. Zhu. 2008. Histone deacetylase inhibitor depsipeptide activates silenced genes through decreasing both CpG and H3K9 methylation on the promoter. *Molecular and cellular biology*. 28:3219-3235.
- Wu, S.Y., A.Y. Lee, H.T. Lai, H. Zhang, and C.M. Chiang. 2013. Phospho switch triggers Brd4 chromatin binding and activator recruitment for gene-specific targeting. *Molecular cell*. 49:843-857.
- Xing, J., D.J. Stewart, J. Gu, C. Lu, M.R. Spitz, and X. Wu. 2008. Expression of methylation-related genes is associated with overall survival in patients with non-small cell lung cancer. *British journal of cancer*. 98:1716-1722.
- Xiong, Y., and K.L. Guan. 2012. Mechanistic insights into the regulation of metabolic enzymes by acetylation. *The Journal of cell biology*. 198:155-164.
- Yamane, K., C. Toumazou, Y. Tsukada, H. Erdjument-Bromage, P. Tempst, J. Wong, and Y. Zhang. 2006. JHDM2A, a JmJC-containing H3K9 demethylase, facilitates transcription activation by androgen receptor. *Cell*. 125:483-495.
- Yamauchi, T., N. Yamauchi, K. Ueki, T. Sugiyama, H. Waki, H. Miki, K. Tobe, S. Matsuda, T. Tsushima, T. Yamamoto, T. Fujita, Y. Taketani, M. Fukayama, S. Kimura, Y. Yazaki, R. Nagai, and T. Kadowaki. 2000. Constitutive tyrosine

RESUMÉ • Chacune de nos cellules possède l'information génétique nécessaire à la constitution de l'organisme entier, mais les cellules différenciées n'expriment qu'un répertoire restreint de gènes. Le contrôle des expressions des gènes apparaît fondamental dans la mise en place et le maintien de l'identité cellulaire. Le contrôle transcriptionnel, premier niveau de régulation de l'expression des gènes, est basé sur l'intégrité de la séquence génique et sur son accessibilité, elle-même conditionnée par les mécanismes épigénétiques contrôlant la dynamique chromatinienne. Dans un contexte pathologique, des altérations génétiques et/ou épigénétiques pourront alors être à l'origine de dérégulations géniques et mener à des fonctions cellulaires altérées. • Lors du développement d'un carcinome bronchique, les cellules pulmonaires acquièrent une capacité de prolifération incontrôlée et une résistance accrue à l'apoptose. Ces caractéristiques phénotypiques, bases de la croissance tumorale, sont le reflet des anomalies qui s'accumulent dans le génome des cellules cancéreuses. Aux altérations génétiques somatiques (mutations ponctuelles, remaniements chromosomiques) s'ajoute un bouleversement global du paysage épigénétique, le tout conduisant une crise de l'identité cellulaire et à des dérégulations massives des expressions géniques. Alors que la répression aberrante de gènes (notamment des gènes suppresseurs de tumeur) a été largement étudiée, l'activation ectopique de gènes normalement silencieux demeure un aspect plus méconnu. Notre hypothèse de travail est que les expressions hors-contexte de gènes tissu-spécifiques puissent être non seulement impliquées dans le phénomène de cancérogenèse, mais également constituer des biomarqueurs tumoraux ou cibles thérapeutiques innovantes. • Dans cette thèse, notre attention s'est portée sur le gène *PRL* codant la prolactine, normalement exprimée dans l'hypophyse et absente du poumon non tumoral. Nous avons détecté une activation ectopique du gène *PRL* dans 10% des tumeurs pulmonaires, principalement des tumeurs neuroendocrines. Nous avons observé que l'expression de *PRL* est associée à des tumeurs particulièrement agressives et à un pronostic sombre pour les patients. Nous avons également démontré que l'expression de *PRL* confère aux cellules cancéreuses pulmonaires une résistance accrue à un stress génotoxique. De manière inattendue, nos résultats suggèrent que l'effet oncogène de l'expression de *PRL* ne repose pas sur les mécanismes d'action classique de la prolactine et nous avons dû remettre en question l'hypothèse initiale d'une sécrétion de l'hormone par les cellules cancéreuses pulmonaires et d'une action autocrine/paracrine au sein de la tumeur via l'activation du récepteur à la prolactine. Le récepteur est en effet absent des cellules cancéreuses pulmonaires et le transcrit *PRL* exprimé est dépourvu de ses premiers exons, ce qui pourrait conduire à la production d'une protéine tronquée de son peptide signal, incapable d'emprunter la voie de sécrétion classique, et par conséquent retenue à l'intérieur de la cellule cancéreuse pulmonaire. Même si les mécanismes d'action *in cellulo* de la prolactine restent à décrypter, nos données suggèrent que l'expression ectopique de *PRL* pourrait constituer une nouvelle cible thérapeutique dans le traitement des tumeurs pulmonaires agressives. • Ce travail de thèse, incluant également des résultats complémentaires obtenus sur trois gènes spécifiques du testicule exprimés de manière aberrante dans les tumeurs pulmonaires (BRDT, SOX30 et SPATA22) met en lumière l'intérêt des expressions ectopiques. Celles-ci peuvent en effet fournir de nouveaux outils de diagnostic et de pronostic aux cliniciens, mais également de nouvelles approches ciblées en complément des thérapies classiques qui ne suffisent pas à limiter la mortalité importante due aux néoplasies pulmonaires.

ABSTRACT • Each human cell contains a genome carrying all the genetic information necessary for the constitution of the whole organism. However, differentiated cells express only a restricted repertoire of genes. The control of gene expressions is fundamental for the establishment and maintenance of cell identity. The first level of gene expression regulation, the transcriptional control, is based on the integrity of the gene sequence but also on its accessibility, itself dependent on a set of epigenetic mechanisms that control chromatin dynamics. In a pathological context, genetic and epigenetic alterations can lead to gene deregulations and altered cellular functions. • During the bronchial carcinogenesis, lung cancer cells acquire a capacity of uncontrolled proliferation and an increased resistance to cell death. These phenotypic characteristics, favoring tumor growth, result from abnormalities that accumulate in the genome of cancer cells. These are somatic genetic alterations, from point mutations to large-scale chromosomal rearrangements, but also a global disruption of the epigenetic landscape – both leading to an identity crisis and to gene deregulations. While the phenomenon of aberrant gene repression (including repression of tumor suppressor genes) has been extensively studied, ectopic activation of normally silent genes remains poorly understood. Our hypothesis is that the “out of context” expression of tissue-specific genes not only could be involved in carcinogenesis, but could also be of high interest as tumor biomarkers or novel therapeutic targets. • In this work, we focused on the prolactin-encoding *PRL* gene, normally mainly expressed in the pituitary gland and absent from non-tumor lung. We detected an ectopic *PRL* gene activation in 10% of lung tumors, mainly neuroendocrine tumors. We observed that *PRL* expression is associated with aggressive tumors and a poor prognosis for patients. We also found that the expression of *PRL* is associated with an increased resistance of lung cancer cells to a genotoxic stress. Unexpectedly, our data suggest that the oncogenic action of *PRL* expression is not based on the conventional mechanisms of prolactin action, and we did not confirm the initial hypothesis of a secretion by lung cancer cells of the prolactin hormone, and its action in an autocrine/paracrine loop within the tumor through the activation of the prolactin receptor. Indeed, the receptor is absent in lung cancer cells and the transcribed *PRL* mRNA is missing its first exons, possibly leading to the production of a truncated prolactin protein, without a functional signal peptide, therefore unable to follow the classical secretion pathway and retained inside the cancer cell. Although the detailed mechanisms of prolactin action in lung cancer remain to be deciphered, our study suggests that the ectopic expression of *PRL* could be used as a new therapeutic target in the treatment of aggressive lung tumors. • This work, also including additional results on three testis-specific genes aberrantly expressed in lung tumors (BRDT, SOX30 and SPATA22) highlights the interest of studying ectopic gene expressions in tumor cells, which can provide new diagnosis and prognosis tools for clinicians as well as new targeted approaches that could be used in addition to conventional lung cancer therapies, which are presently insufficient to limit the high mortality due to lung neoplasms.